(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 0 880 590 B1

(12) **EUROPEAN PATENT SPECIFICATION**

(45) Date of publication and mention of the grant of the patent:
20.10.2004 Bulletin 2004/43

(21) Application number: **97906528.1**

(22) Date of filing: **11.02.1997**

(51) Int Cl.⁷: **C12N 15/55**, C12N 9/16, C12N 1/21, C12P 13/02, C12Q 1/68, C07K 16/40, G01N 33/53

(86) International application number:
PCT/US1997/002039

(87) International publication number:
WO 1997/030160 (21.08.1997 Gazette 1997/36)

(54) ESTERASES
ESTERASEN
ESTERASES

(84) Designated Contracting States:
AT BE CH DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE

(30) Priority: **16.02.1996 US 602359**

(43) Date of publication of application:
02.12.1998 Bulletin 1998/49

(60) Divisional application:
04017203.3

(73) Proprietor: **DIVERSA CORPORATION**
San Diego, California 92121 (US)

(72) Inventors:
- **ROBERTSON, Dan, E.**
  Haddonfield, NJ 08033 (US)
- **MURPHY, Dennis**
  Paoli, PA 19301 (US)
- **REID, John**
  Bryn Mawr, PA 19010 (US)
- **MAFFIA, Anthony, M.**
  Wilmington, DE 19810 (US)
- **LINK, Steven**
  Wilmington, DE 19803 (US)
- **SWANSON, Ronald, V.**
  Media, PA 19063 (US)
- **WARREN, Patrick, V.**
  Philadelphia, PA 19163 (US)
- **KOSMOTKA, Anna**
  Brookhaven, PA 19015 (US)
- **CALLEN, Walter**
  San Diego, CA 92117 (US)

(74) Representative: **VOSSIUS & PARTNER**
Siebertstrasse 4
81675 München (DE)

(56) References cited:

EP-A- 0 318 775	**WO-A-93/25691**
WO-A-94/02617	**WO-A-95/23853**
WO-A-95/34644	**US-A- 5 306 636**

- **WRAY L. ET AL.: "Streptomyces coelicolor transcription factor, AC L03213" EBI DATABASE,9 October 1992 (1992-10-09), XP002145895**
- **KUGIMIYA WATARU ET AL: "Molecular cloning and structure of the gene for esterase from a thermophilic bacterium, Bacillus stearothermophilus IFO 12550." BIOSCIENCE BIOTECHNOLOGY AND BIOCHEMISTRY 1992, vol. 56, no. 12, 1992, pages 2074-2075, XP002138593 ISSN: 0916-8451**
- **HUDDLESTON S ET AL: "The identification and partial characterisation of a novel inducible extracellular thermostable esterase from the archaeon Sulfolobus shibatae." BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS 1995, vol. 216, no. 2, 1995, pages 495-500, XP002138594 ISSN: 0006-291X**

158

Note: Within nine months from the publication of the mention of the grant of the European patent, any person may give notice to the European Patent Office of opposition to the European patent granted. Notice of opposition shall be filed in a written reasoned statement. It shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

Printed by Jouve, 75001 PARIS (FR)

(Cont. next page)

- DERWENT BIOTECHNOLOGY ABSTRACTS, Vol. 14, No. 15, 17 July 1995, page 125, Abstract No. 95-09223, LADRAT C. et al., "Thermostable Enzymes Screened on Thermophilic Microorganisms from Deep-Sea Hydrothermal Vents - -beta- Glucosidase, Alcohol Dehydrogenase, Protease and Esterase Activity"; & C.R. SEANCES ACAD. SCI. SER 3., 1995, 318(4), 423-429.
- DERWENT BIOTECHNOLOGY ABSTRACTS, Vol. 13, No. 15, 25 July 1994, pages 95-96, Abstract No. 94-08869, CORNEC L. et al., "Thermostable Esterase Activity from Deep-Sea Sulfur Dependent Thermophilic Archaebacteria-Thermostable Enzyme Purification from Thermophilic Archaebacterium and Characterization"; & THERMOPHILES SCI. TECHNOL., 1992, 18.
- APPL. MICROBIOL. BIOTECHNOL., 1994, Vol. 42, CANGANELLA F., "Characterization of Amylolytic and Pullulytic Enzymes from Thermophilic Archea and from a New Fervidobacterium Species", pages 239-245.

Description

[0001] This invention relates to newly identified polynucleotides, polypeptides encoded by such polynucleotides. the use of such polynucleotides and polypeptides, as well as the production and isolation of such polynucleotides and polypeptides. More particularly. the polynucleotides and polypeptides of the present invention have been putatively identified as esterases. Esterases are enzymes that catalyze the hydrolysis of ester groups to organic acids and alcohols.

[0002] Many esterases are known and have been discovered in a broad variety of organisms. including bacteria. yeast and higher animals and plants. A principal example of esterases are the lipases. which are used in the hydrolysis of lipids, acidolysis(replacement of an esterified fatty acid with a free fatty acid) reactions, transesterification(exchange of fatty acids between triglycerides)reactions, and in ester synthesis. The major industrial applications for lipases include: the detergent industry, where they are employed to decompose fatty materials in laundry stains into easily removable hydrophilic substances: the food and beverage industry where they are used in the manufacture of cheese, the ripening and flavoring of cheese, as antistaling agents for bakery products. and in the production of margarine and other spreads with natural butter flavors; in waste systems; and in the pharmaceutical industry where they are used as digestive aids.

[0003] The polynucleotides and polypeptides of the present invention have been identified as esterases as a result of their enzymatic activity.

[0004] In accordance with one aspect of the present invention, there are provided novel enzymes. as well as active fragments, analogs and derivatives thereof.

[0005] In accordance with another aspect of the present invention, there are provided isolated nucleic acid molecules encoding the enzymes of the present invention including mRNAs, cDNAs, genomic DNAs as well as active analogs and fragments of such enzymes.

[0006] In accordance with yet a further aspect of the present invention, there is provided a process for producing such polypeptides by recombinant techniques comprising culturing recombinant prokaryotic and/or eukaryotic host cells, containing a nucleic acid sequence of the present invention, under conditions promoting expression of said enzymes and subsequent recovery of said enzymes.

[0007] In accordance with yet a further aspect of the present invention, there is provided a process for utilizing such enzymes. or polynucleotides encoding such enzymes for hydrolyzing ester groups to yield an organic acid and an alcohol. The esterases of the invention are stable at high temperatures and in organic solvents and, thus, are superior for use in production of optically pure chiral compounds used in pharmaceutical, agricultural and other chemical industries.

[0008] In accordance with yet a further aspect of the present invention, there are also provided nucleic acid probes comprising nucleic acid molecules of sufficient length to hybridize to a nucleic acid sequence of the present invention.

[0009] In accordance with yet a further aspect of the present invention, there is provided a process for utilizing such enzymes, or polynucleotides encoding such enzymes, for *in vitro* purposes related to scientific research, for example, to generate probes for identifying similar sequences which might encode similar enzymes from other organisms by using certain regions, *i.e.*, conserved sequence regions, of the nucleotide sequence.

[0010] These and other aspects of the present invention should be apparent to those skilled in the art from the teachings herein.

[0011] The following drawings are illustrative of embodiments of the invention and are not meant to limit the scope of the invention as encompassed by the claims.

[0012] Figure 1 is an illustration of the full-length DNA (SEQ ID NO:23) and corresponding deduced amino acid sequence (SEQ ID NO:33) of *Staphylothermus marinus* F1-12LC of the present invention. Sequencing was performed using a 378 automated DNA sequencer (Applied Biosystems, Inc.) for all sequences of the present invention.

[0013] Figure 2 is an illustration of the full-length DNA (SEQ ID NO:24) and corresponding deduced amino acid sequence (SEQ ID NO:34) of *Pyrodictium* TAG 11-17LC.

[0014] Figure 3 is an illustration of the full-length DNA (SEQ ID NO:25) and corresponding deduced amino acid sequence (SEQ ID NO:35) of *Archaeoglobus venificus* SNP6-24LC.

[0015] Figure 4 is an illustration of the full-length DNA (SEQ ID NO:26) and corresponding deduced amino acid sequence (SEQ ID NO:36) of *Aquifex pyrophilus*-28LC.

[0016] Figure 5 is an illustration of the full-length DNA (SEQ ID NO:27) and corresponding deduced amino acid sequence (SEQ ID NO:37) of M11TL-29L.

[0017] Figure 6 is an illustration of the full-length DNA (SEQ ID NO:28) and corresponding deduced amino acid sequence (SEQ ID NO:38) of *Thermococcus* CL-2-30LC.

[0018] Figure 7 is an illustration of the full-length DNA (SEQ ID NO:29) and corresponding deduced amino acid sequence (SEQ ID NO:39) of *Aquifex* VF5-34LC.

[0019] Figure 8 is an illustration of the full-length DNA (SEQ ID NO:30) and corresponding deduced amino acid sequence (SEQ ID NO:40) of *Teredinibacter*-42L.

[0020] Figure 9 is an illustration of the full-length DNA (SEQ ID NO:31) and corresponding deduced amino acid sequence (SEQ ID NO:41) of *Archaeoglobus fulgidus VC*16-16MC.

[0021] Figure 10 is an illustration of the full-length DNA (SEQ ID NO:32) and corresponding deduced amino acid sequence (SEQ ID NO:42) of *Sulfolobus solfataricus* P1-8LC.

[0022] Figure 11 is an illustration of the full-length DNA (SEQ ID NO:43) and corresponding deduced amino acid sequence (SEQ ID NO:53) of LA11.1 Esterase es2 of the present invention.

[0023] Figure 12 is an illustration of the full-length DNA (SEQ ID NO:44) and corresponding deduced amino acid sequence (SEQ ID NO:54) of Whale Mat Sample 11.801 Esterase es9.

[0024] Figure 13 is an illustration of the full-length DNA (SEQ ID NO:45) and corresponding deduced amino acid sequence (SEQ ID NO:55) of *Metallosphaera Prunae Ron* 12/2 Esterase 23mc1 .

[0025] Figure 14 is an illustration of the full-length DNA (SEQ ID NO:46) and corresponding deduced amino acid sequence (SEQ ID NO:56) of *Thermotoga. neapolitana* 5068 Esterase 56mc4.

[0026] Figure 15 is an illustration of the full-length DNA (SEQ ID NO:47) and corresponding deduced amino acid sequence (SEQ ID NO:57) of *Melittangium lichenicola* Esterase 77mc1.

[0027] Figure 16 is an illustration of the full-length DNA (SEQ ID NO:48) and corresponding deduced amino acid sequence (SEQ ID NO:58) of Whale Mat Sample 11.801 Esterase es2.

[0028] Figure 17 is an illustration of the full-length DNA (SEQ ID NO:49) and corresponding deduced amino acid sequence (SEQ ID NO:59) of Whale Mat Sample AD3059 Esterase es4.

[0029] Figure 18 is an illustration of the full-length DNA (SEQ ID NO:50) and corresponding deduced amino acid sequence (SEQ ID NO:60) of *Microscilla furvescens* Esterase 53sc2.

[0030] figure 19 is an illustration of the full-length DNA (SEQ ID NO:51) and corresponding deduced amino acid sequence (SEQ ID NO:61) of *Thermotoga maritima MSB8* Esterase 6sc1.

[0031] Figure 20 is an illustration of the full-length DNA (SEQ ID NO:52) and corresponding deduced amino acid sequence (SEQ ID NO:62) of *Polyangium brachysporum* Esterase 78mc1.

[0032] The term "gene" means the segment of DNA involved in producing a polypeptide chain; it includes regions preceding and following the coding region (leader and trailer) as well as intervening sequences (introns) between individual coding segments (exons).

[0033] A coding sequence is "operably linked to" another coding sequence when RNA polymerase will transcribe the two coding sequences into a single mRNA, which is then translated into a single polypeptide having amino acids derived from both coding sequences. The coding sequences need not be contiguous to one another so long as the expressed sequences ultimately process to produce the desired protein.

[0034] "Recombinant" enzymes refer to enzymes produced by recombinant DNA techniques; *i.e.*, produced from cells transformed by an exogenous DNA construct encoding the desired enzyme. "Synthetic" enzymes are those prepared by chemical synthesis.

[0035] A DNA "coding sequence of" or a "nucleotide sequence encoding" a particular enzyme, is a DNA sequence which is transcribed and translated into an enzyme when placed under the control of appropriate regulatory sequences.

[0036] In accordance with an aspect of the present invention, there are provided isolated nucleic acids (polynucleotides) which encode for the mature enzymes having the deduced amino acid sequences of Figures 1-20 (SEQ ID NOS:23-32 and 43 to 52.)

[0037] The deposit(s) have been made under the terms of the Budapest Treaty on the International Recognition of the deposit of micro-organisms for purposes of patent procedure. The strains will be irrevocably and without restriction or condition released to the public upon the issuance of a patent. These deposits are provided merely as convenience to those of skill in the art and are not an admission that a deposit would be required under 35 U.S.C. §112. The sequences of the polynucleotides contained in the deposited materials, as well as the amino acid sequences of the polypeptides encoded thereby, are controlling in the event of any conflict with any description of sequences herein. A license may be required to make, use or sell the deposited materials, and no such license is hereby granted.

[0038] The polynucleotides of this invention were originally recovered from genomic gene libraries derived from the following organisms:

[0039] *Staphylothermus marinus* F1 is a thermophilic sulfur archaea which was isolated in Vulcano, Italy. It grows optimally at 85°C (T_{max} = 98°C) at pH 6.5.

[0040] *Pyrodictium* TAG11 is a thermophilic sulfur archaea which was isolated in the Middle Atlantic Ridge. It grows optimally at 103°C (T_{max} = 110°C) at pH 6.5.

[0041] *Archaeoglobus venificus* SNP6 was isolated in the Middle Atlantic Ridge and grows optimally at 75°C (T_{max} = 92°C) at pH 6.9.

[0042] *Aquifex pyrophilus* K01 5a was isolated at Kolbeinsey Ridge, North of Iceland. This marine organism is a gram-negative, rod-shaped, strictly chemolithoautrophic, knall gas bacterium. It grows optimally at 85°C (T_{max} = 95°C) at pH 6.8.

[0043] M11TL is a new species of Desulfurococcus which was isolated from Diamond Pool (formerly Jim's Black Pool) in Yellowstone. The organism grows heterotrophically by fermentation of different organic materials (sulfur is not necessary) in grape-like aggregates optimally at 85 - 88°C in a low salt medium at pH 7.0.

[0044] *Thermococcus* CL-2 was isolated in the North Cleft Segment of the Juan de Fuca Ridge from a severed alvinellid worm residing on a "black smoker" sulfide structure. This marine archaea forms pleomorphic cocci, and grows optimally at 88°C.

[0045] *Aquifex* VF5 was isolated at a beach in Vulcano, Italy. This marine organism is a gram-negative, rod-shaped, strictly chemolithoautotrophic, knall gas bacterium. It grows optimally at 85°C (T_{max} = 95°C) at pH 6.8.

[0046] *Teredinibacter* (pure) is an endosymbiont of the shipworm *Bankia gouldi*. The organism has straight to slightly bent 5-10 µm rods, and forms spiral cells as stationary phase is met. The organism was described in Science (1983) 22:1401-1403. It grows optimally at 30°C at pH 8.0.

[0047] *Archaeoglobus fulgidus VC*16 was isolated in Vulcano, Italy. The organism grows optimally at 85°C (T_{max} = 92°C) at pH 7.0.

[0048] *Sulfolobus solfataricus* P1 grows optimally at 85°C (T_{max} = 87°C) at pH 2.0.

[0049] Accordingly, the polynucleotides and enzymes encoded thereby are identified by the organism from which they were isolated. and are sometimes hereinafter referred to as F1/12LC (Figure 1 and SEQ ID NOS:23 and 33), TAG11/17LC (Figure 2 and SEQ ID NOS:24 and 34), SNP6/24LC (Figure 3 and SEQ ID NOS:25 and 35), AqP/28LC (Figure 4 and SEQ ID NOS:26 and 36), M11TL/29L (Figure 5 and SEQ ID NOS:27 and 37), CL-2/30LC (Figure 6 and SEQ ID NOS:28 and 38), VF5/34LC (Figure 7 and SEQ ID NOS:29 and 39), Trb/42L (Figure 8 and SEQ ID NOS:30 and 40). VC16/16MC (Figure 9 and SEQ ID NOS:31 and 41) and P1/8LC (Figure 10 and SEQ ID NOS: 32 and 42).

[0050] The polynucleotides and polypeptides of the present invention show identity at the nucleotide and protein level to known genes and proteins encoded thereby as shown in Table 1.

Table 1

Enzyme	Gene w/closest Homology (Organism)	Protein Similarity (%)	Protein Identity (%)	DNA Identity (%)
F1/12LC	No significant homology	-	-	-
TAG11/17LC	No significant homology	-	-	-
SNP6/24LC	PIR S34609 - carboxylesterase *Pseudomones sp.* (strain KWI-56) open reading frame of unknown function in *E.coli.*	46	27	42
AqP/29LC		53	31	38
M11TL/29LC	No significant homology	-	-	-
CL02/30LC	No significant homology	-	-	-
VF5/34LC	Identified by homology to 28LC; also homologous to ORF of unknown function 5' of tgs in *E. coli*	84	71	71
Trb/42L	No significant homology	-	-	-
P1-8LC				
VC16-16MC				

[0051] All the clones identified in Table 1 encode polypeptides which have esterase activity.

[0052] This invention, in addition to the isolated nucleic acid molecules encoding the enzymes of the present invention. also provides substantially similar sequences. Isolated nucleic acid sequences are substantially similar if: (i) they are capable of hybridizing under conditions hereinafter described, to the polynucleotides of SEQ ID NOS:23-32 and 43-52; (ii) or they encode DNA sequences which are degenerate to the polynucleotides of SEQ ID NOS:23-32 and 43 to 52 Degenerate DNA sequences encode the amino acid sequences of SEQ ID NOS:33-42 and 53 to 62, but have variations in the nucleotide coding sequences. As used herein, substantially similar refers to the sequences having similar identity to the sequences of the instant invention. The nucleotide sequences that are substantially the same can be identified by hybridization or by sequence comparison. Enzyme sequences that are substantially the same can be identified by one or more of the following: proteolytic digestion, gel electrophoresis and/or microsequencing.

[0053] One means for isolating the nucleic acid molecules encoding the enzymes of the present invention is to probe a gene library with a natural or artificially designed probe using art recognized procedures (see, for example: Current Protocols in Molecular Biology, Ausubel F.M. *et al.* (EDS.) Green Publishing Company Assoc. and John Wiley Interscience, New York, 1989, 1992). It is appreciated by one skilled in the art that the polynucleotides of SEQ ID NOS: 23-32 and 43 to 52 or fragments thereof (comprising at least 12 contiguous nucleotides), are particularly useful probes. Other particularly useful probes for this purpose are hybridizable fragments of the sequences of SEQ ID NOS:1-22 (*i. e.*, comprising at least 12 contiguous nucleotides).

[0054] With respect to nucleic acid sequences which hybridize to specific nucleic acid sequences disclosed herein, hybridization may be carried out under conditions of reduced stringency, medium stringency or even stringent conditions. As an example of oligonucleotide hybridization. a polymer membrane containing immobilized denatured nucleic acids is first prehybridized for 30 minutes at 45°C in a solution consisting of 0.9 M NaCl, 50 mM NaH_2PO_4, pH 7.0. 5.0 mM Na_2EDTA, 0.5% SDS, 10X Denhardt's, and 0.5 mg/mL polyriboadenylic acid. Approximately 2 X 10^7 cpm (specific activity 4-9 X 10^8 cpm/ug) of ^{32}P end-labeled oligonucleotide probe are then added to the solution. After 12-16 hours of incubation, the membrane is washed for 30 minutes at room temperature in IX SET (150 mM NaCl, 20 mM Tris hydrochloride, pH 7.8, 1 mM Na_2EDTA) containing 0.5% SDS. followed by a 30 minute wash in fresh 1X SET at Tm -10°C for the oligo-nucleotide probe. The membrane is then exposed to autoradiographic film for detection of hybridization signals.

[0055] Stringent conditions means hybridization will occur only if there is at least 90% identity, preferably at least 95% identity and most preferably at least 97% identity between the sequences. *See* J. Sambrook *et al., Molecular Cloning, A Laboratory Manual, 2d Ed.*, Cold Spring Harbor Laboratory (1989) which is hereby incorporated by reference in its entirety.

[0056] As used herein, a first DNA (RNA) sequence is at least 70% and preferably at least 80% identical to another DNA (RNA) sequence if there is at least 70% and preferably at lest a 80% or 90% identity, respectively, between the bases of the first sequence and the bases of the another sequence, when property aligned with each other. for example when aligned by BLASTN.

[0057] The present invention relates to polynucleotides which differ from the reference polynucleotide such that the changes are silent changes, for example the change do not alter the amino acid sequence encoded by the polynucleotide. The present invention also relates to nucleotide changes which result in amino acid substitutions, additions, deletions, fusions and truncations in the polypeptide encoded by the reference polynucleotide. In a preferred aspect of the invention these polypeptides retain the same biological action as the polypeptide encoded by the reference polynucleotide.

[0058] The polynucleotides of this invention were recovered from genomic gene libraries from the organisms listed in Table 1. Gene libraries were generated in the Lambda ZAP II cloning vector (Stratagene Cloning Systems). Mass excisions were performed on these libraries to generate libraries in the pBluescript phagemid. Libraries were generated and excisions were performed according to the protocols/methods hereinafter described.

[0059] The polynucleotides of the present invention may be in the form of RNA or DNA which DNA includes cDNA, genomic DNA, and synthetic DNA. The DNA may be double-stranded or single-stranded, and if single stranded may be the coding strand or non-coding (anti-sense) strand. The coding sequences which encodes the mature enzymes may be identical to the coding sequences shown in Figures 1-20 (SEQ ID NOS:23-32) and 43 to 52 or may be a different coding sequence which coding sequence, as a result of the redundancy or degeneracy of the genetic code, and 43-52, encodes the same mature enzymes as the DNA of Figures 1-20 (SEQ ID NOS:23-32 and 43-52).

[0060] The polynucleotide which encodes for the mature enzyme of Figures 1-20 (SEQ ID NOS:33-42) and 53-62 may include, but is not limited to: only the coding sequence for the mature enzyme; the coding sequence for the mature enzyme and additional coding sequence such as a leader sequence or a proprotein sequence; the coding sequence for the mature enzyme (and optionally additional coding sequence) and non-coding sequence, such as introns or non-coding sequence 5' and/or 3' of the coding sequence for the mature enzyme.

[0061] Thus, the term "polynucleotide encoding an enzyme (protein)" encompasses a polynucleotide which includes only coding sequence for the enzyme as well as a polynucleotide which includes additional coding and/or non-coding sequence.

[0062] The present invention further relates to variants of the hereinabove described polynucleotides which encode for fragments, analogs and derivatives of the enzymes having the deduced amino acid sequences of Figures 1-20 (SEQ ID NOS:33-42 and 53-62). The variant of the polynucleotide may be a naturally occurring allelic variant of the polynucleotide or a non-naturally occurring variant of the polynucleotide.

[0063] Thus. the present invention includes polynucleotides encoding the same mature enzymes as shown in Figures 1-20 (SEQ ID NOS:23-32 and 43-52) as well as variants of such polynucleotides which variants encode for a fragment, derivative or analog of the enzymes of Figures 1-20 (SEQ ID NOS:23-32 and 43 to 52). Such nucleotide variants include deletion variants, substitution variants and addition or insertion variants.

[0064] As hereinabove indicated, the polynucleotides may have a coding sequence which is a naturally occurring allelic variant of the coding sequences shown in Figures 1-20 (SEQ ID NOS:23-32 and 43 to 52). As known in the art, an allelic variant is an alternate form of a polynucleotide sequence which may have a substitution, deletion or addition of one or more nucleotides, which does not substantially alter the function of the encoded enzyme.

[0065] Fragments of the full length gene of the present invention may be used as hybridization probes for a cDNA or a genomic library to isolate the full length DNA and to isolate other DNAs which have a high sequence similarity to the gene or similar biological activity. Probes of this type preferably have at least 10, preferably at least 15, and even more preferably at least 30 bases and may contain, for example, at least 50 or more bases. The probe may also be used to identify a DNA clone corresponding to a full length transcript and a genomic clone or clones that contain the complete gene including regulatory and promoter regions, exons and introns. An example of a screen comprises isolating the coding region of the gene by using the known DNA sequence to synthesize an oligonucleotide probe. Labeled oligonucleotides having a sequence complementary to that of the gene of the present invention are used to screen a library of genomic DNA to determine which members of the library the probe hybridizes to.

[0066] It is also appreciated that such probes can be and are preferably labeled with an analytically detectable reagent to facilitate identification of the probe. Useful reagents include but are not limited to radioactivity, fluorescent dyes or enzymes capable of catalyzing the formation of a detectable product. The probes are thus useful to isolate complementary copies of DNA from other sources or to screen such sources for related sequences.

[0067] The present invention further relates to polynucleotides or fragments thereof encoding an enzyme having esterase activity and which hybridize to the hereinabove-described sequences if there is at least 70%, preferably at least 90%, and more preferably at least 95% identity between the sequences. The present invention particularly relates to polynucleotides which hybridize under stringent conditions to the hereinabove-described polynucleotides. As herein used, the term "stringent conditions" means hybridization will occur only if there is at least 95 % and preferably at least 97 % identity between the sequences. The polynucleotides which hybridize to the hereinabove described polynucleotides in a preferred embodiment encode enzymes which either retain substantially the same biological function or activity as the mature enzyme encoded by the DNA of Figures 1-20 (SEQ ID NOS:23-32 and 43 to 52.)

[0068] Alternatively, the polynucleotide may have at least 15 bases, preferably at least 30 bases, and more preferably at least 50 bases of a polynucleotide of the present invention and which has an identity thereto, as hereinabove described, and which may or may not retain activity. For example, such polynucleotides may be employed as probes for the polynucleotides of SEQ ID NOS:23-32, and 43 to 52 for example, for recovery of the polynucleotide or as a diagnostic probe or as a PCR primer.

[0069] Thus, the present invention is directed to polynucleotides or fragments thereof encoding an enzyme having esterase activity and having at least a 70% identity, preferably at least 90% identity and more preferably at least a 95% identity to a polynucleotide which encodes the enzymes of SEQ ID NOS:33-42 and 53 to 62 as well as fragments thereof, which fragments have at least 15 bases, preferably at least 30 bases and most preferably at least 50 bases. which fragments are at least 90% identical, preferably at least 95 % identical and most preferably at least 97% identical under stringent conditions to any portion of a polynucleotide of the present invention.

[0070] The present invention further relates to enzymes which have the deduced amino acid sequences of Figures 1-20 (SEQ ID NOS:23-32 and 43 to 52) as well as fragments, analogs and derivatives of such enzyme.

[0071] The terms "fragment," "derivative" and "analog" when referring to the enzymes of Figures 1-20 (SEQ ID NOS: 33-42 and 53 to 62) mean enzymes which retain essentially the same biological function or activity as such enzymes. Thus, an analog includes a proprotein which can be activated by cleavage of the proprotein portion to produce an active mature enzyme.

[0072] The enzymes of the present invention may be a recombinant enzyme, a natural enzyme or a synthetic enzyme, preferably a recombinant enzyme.

[0073] The fragment, derivative or analog of the enzymes of Figures 1-20 (SEQ ID NOS:33-42 and 53 to 62) may be (i) one in which one or more of the amino acid residues are substituted with a conserved or non-conserved amino acid residue (preferably a conserved amino acid residue) and such substituted amino acid residue may or may not be one encoded by the genetic code. or (ii) one in which one or more of the amino acid residues includes a substituent group, or (iii) one in which the mature enzyme is fused with another compound, such as a compound to increase the half-life of the enzyme (for example, polyethylene glycol), or (iv) one in which the additional amino acids are fused to the mature enzyme, such as a leader or secretory sequence or a sequence which is employed for purification of the mature enzyme or a proprotein sequence. Such fragments, derivatives and analogs are deemed to be within the scope of those skilled in the art from the teachings herein.

[0074] The enzymes and polynucleotides of the present invention are preferably provided in an isolated form, and preferably are purified to homogeneity.

[0075] The term "isolated" means that the material is removed from its original environment (e.g., the natural environment if it is naturally occurring). For example, a naturally-occurring polynucleotide or enzyme present in a living animal is not isolated, but the same polynucleotide or enzyme, separated from some or all of the coexisting materials in the natural system, is isolated. Such polynucleotides could be part of a vector and/or such polynucleotides or enzymes could be part of a composition, and still be isolated in that such vector or composition is not part of its natural environment.

[0076] The enzymes of the present invention include the enzymes of SEQ ID NOS:33-42 and 53 to 62 (in particular the mature enzyme) as well as enzymes which have at least 70 % similarity (preferably at least 70% identity) to the enzymes of SEQ ID NOS:33-42 and 53 to 62 and more preferably at least 90% similarity (more preferably at least 90% identity) to the enzymes of SEQ ID NOS:33-42 and 53 to 62 and still more preferably at least 95% similarity (still more preferably at least 95% identity) to the enzymes of SEQ ID NOS:33-42 and 53 to 62 and also include portions of such enzymes with such portion of the enzyme generally containing at least 30 amino acids and more preferably at least 50 amino acids.

[0077] As known in the art "similarity" between two enzymes is determined by comparing the amino acid sequence and its conserved amino acid substitutes of one enzyme to the sequence of a second enzyme.

[0078] A variant, *i.e.* a "fragment", "analog" or "derivative" polypeptide, and reference polypeptide may differ in amino acid sequence by one or more substitutions, additions, deletions, fusions and truncations, which may be present in any combination.

[0079] Among preferred variants are those that vary from a reference by conservative amino acid substitutions. Such substitutions are those that substitute a given amino acid in a polypeptide by another amino acid of like characteristics. Typically seen as conservative substitutions are the replacements, one for another, among the aliphatic amino acids Ala, Val, Leu and Ile; interchange of the hydroxyl residues Ser and Thr, exchange of the acidic residues Asp and Glu, substitution between the amide residues Asn and Gln, exchange of the basic residues Lys and Arg and replacements among the aromatic residues Phe, Tyr.

[0080] Most highly preferred are variants which retain the same biological function and activity as the reference polypeptide from which it varies.

[0081] Fragments or portions of the enzymes of the present invention may be employed for producing the corresponding full-length enzyme by peptide synthesis; therefore, the fragments may be employed as intermediates for producing the full-length enzymes. Fragments or portions of the polynucleotides of the present invention may be used to synthesize full-length polynucleotides of the present invention.

[0082] The present invention also relates to vectors which include polynucleotides of the present invention, host cells which are genetically engineered with vectors of the invention and the production of enzymes of the invention by recombinant techniques.

[0083] Host cells are genetically engineered (transduced or transformed or transfected) with the vectors of this invention which may be, for example, a cloning vector or an expression vector. The vector may be, for example, in the form of a plasmid, a viral particle, a phage, *etc.* The engineered host cells can be cultured in conventional nutrient media modified as appropriate for activating promoters, selecting transformants or amplifying the genes of the present invention. The culture conditions, such as temperature, pH and the like, are those previously used with the host cell selected for expression, and will be apparent to the ordinarily skilled artisan.

[0084] The polynucleotides of the present invention may be employed for producing enzymes by recombinant techniques. Thus, for example, the polynucleotide may be included in any one of a variety of expression vectors for expressing an enzyme. Such vectors include chromosomal. nonchromosomal and synthetic DNA sequences, *e.g.*, derivatives of SV40; bacterial plasmids; phage DNA; baculovirus; yeast plasmids; vectors derived from combinations of plasmids and phage DNA, viral DNA such as vaccinia, adenovirus. fowl pox virus, and pseudorabies. However, any other vector may be used as long as it is replicable and viable in the host.

[0085] The appropriate DNA sequence may be inserted into the vector by a variety of procedures. In general. the DNA sequence is inserted into an appropriate restriction endonuclease site(s) by procedures known in the art. Such procedures and others are deemed to be within the scope of those skilled in the art.

[0086] The DNA sequence in the expression vector is operatively linked to an appropriate expression control sequence(s) (promoter) to direct mRNA synthesis. As representative examples of such promoters, there may be mentioned: LTR or SV40 promoter, the *E. coli. lac* or *trp,* the phage lambda P_L promoter and other promoters known to control expression of genes in prokaryotic or eukaryotic cells or their viruses. The expression vector also contains a ribosome binding site for translation initiation and a transcription terminator. The vector may also include appropriate sequences for amplifying expression.

[0087] In addition, the expression vectors preferably contain one or more selectable marker genes to provide a phenotypic trait for selection of transformed host cells such as dihydrofolate reductase or neomycin resistance for eukaryotic cell culture, or such as tetracycline or ampicillin resistance in *E. coli.*

[0088] The vector containing the appropriate DNA sequence as hereinabove described, as well as an appropriate promoter or control sequence, may be employed to transform an appropriate host to permit the host to express the protein.

[0089] As representative examples of appropriate hosts, there may be mentioned: bacterial cells, such as *E. coli.*

Strepromyces, Bacillus subtilis; fungal cells, such as yeast; insect cells such as *Drosophila S2* and *Spodoptera Sf9*; animal cells such as CHO, COS or Bowes melanoma: adenoviruses; plant cells, *etc.* The selection of an appropriate host is deemed to be within the scope of those skilled in the art from the teachings herein.

[0090] More particularly, the present invention also includes recombinant constructs comprising one or more of the sequences as broadly described above. The constructs comprise a vector, such as a plasmid or viral vector, into which a sequence of the invention has been inserted, in a forward or reverse orientation. In a preferred aspect of this embodiment, the construct further comprises regulatory sequences, including, for example, a promoter, operably linked to the sequence. Large numbers of suitable vectors and promoters are known to those of skill in the art, and are commercially available. The following vectors are provided by way of example; Bacterial: pQE70, pQE60, pQE-9 (Qiagen), pBluescript II KS, ptrc99a, pKK223-3, pDR540, pRIT2T (Pharmacia); Eukaryotic: pXT1, pSG5 (Stratagene) pSVK3, pBPV, pMSG, pSVL, SV40 (Pharmacia). However, any other plasmid or vector may be used as long as they are replicable and viable in the host.

[0091] Promoter regions can be selected from any desired gene using CAT (chloramphenicol transferase) vectors or other vectors with selectable markers. Two appropriate vectors are pKK232-8 and pCM7. Particular named bacterial promoters include lacI, lacZ, T3, T7, gpt. lambda P_R, P_L and trp. Eukaryotic promoters include CMV immediate early, HSV thymidine kinase, early and late SV40, LTRs from retrovirus, and mouse metallothionein-I. Selection of the appropriate vector and promoter is well within the level of ordinary skill in the art.

[0092] In a further embodiment, the present invention relates to host cells containing the above-described constructs. The host cell can be a higher eukaryotic cell, such as a mammalian cell, or a lower eukaryotic cell, such as a yeast cell, or the host cell can be a prokaryotic cell, such as a bacterial cell. Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-Dextran mediated transfection, or electroporation (Davis, L., Dibner, M., Battey, I., *Basic Methods in Molecular Biology,* (1986)).

[0093] The constructs in host cells can be used in a conventional manner to produce the gene product encoded by the recombinant sequence. Alternatively, the enzymes of the invention can be synthetically produced by conventional peptide synthesizers.

[0094] Mature proteins can be expressed in mammalian cells, yeast, bacteria, or other cells under the control of appropriate promoters. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention. Appropriate cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described by Sambrook *et al., Molecular Cloning: A Laboratory Manual, Second Edition,* Cold Spring Harbor, N.Y., (1989), the disclosure of which is hereby incorporated by reference.

[0095] Transcription of the DNA encoding the enzymes of the present invention by higher eukaryotes is increased by inserting an enhancer sequence into the vector. Enhancers are cis-acting elements of DNA, usually about from 10 to 300 bp that act on a promoter to increase its transcription. Examples include the SV40 enhancer on the late side of the replication origin bp 100 to 270, a cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers.

[0096] Generally, recombinant expression vectors will include origins of replication and selectable markers permitting transformation of the host cell, *e.g.,* the ampicillin resistance gene of *E. coli* and *S. cerevisiae* TRP1 gene, and a promoter derived from a highly-expressed gene to direct transcription of a downstream structural sequence. Such promoters can be derived from operons encoding glycolytic enzymes such as 3-phosphoglycerate kinase (PGK). α-factor, acid phosphatase, or heat shock proteins, among others. The heterologous structural sequence is assembled in appropriate phase with translation initiation and termination sequences, and preferably, a leader sequence capable of directing secretion of translated enzyme. Optionally, the heterologous sequence can encode a fusion enzyme including an N-terminal identification peptide imparting desired characteristics, e.g., stabilization or simplified purification of expressed recombinant product.

[0097] Useful expression vectors for bacterial use are constructed by inserting a structural DNA sequence encoding a desired protein together with suitable translation initiation and termination signals in operable reading phase with a functional promoter. The vector will comprise one or more phenotypic selectable markers and an origin of replication to ensure maintenance of the vector and to, if desirable, provide amplification within the host. Suitable prokaryotic hosts for transformation include *E. coli, Bacillus subtilis, Salmonella typhimurium* and various species within the genera Pseudomonas, Streptomyces, and Staphylococcus, although others may also be employed as a matter of choice.

[0098] As a representative but nonlimiting example, useful expression vectors for bacterial use can comprise a selectable marker and bacterial origin of replication derived from commercially available plasmids comprising genetic elements of the well known cloning vector pBR322 (ATCC 37017). Such commercial vectors include, for example, pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Sweden) and GEM1 (Promega Biotec, Madison, WI, USA). These pBR322 "backbone" sections are combined with an appropriate promoter and the structural sequence to be expressed.

[0099] Following transformation of a suitable host strain and growth of the host strain to an appropriate cell density, the selected promoter is induced by appropriate means (*e.g.,* temperature shift or chemical induction) and cells are cultured for an additional period.

[0100] Cells are typically harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification.

[0101] Microbial cells employed in expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption. or use of cell lysing agents, such methods are well known to those skilled in the art.

[0102] Various mammalian cell culture systems can also be employed to express recombinant protein. Examples of mammalian expression systems include the COS-7 lines of monkey kidney fibroblasts, described by Gluzman, *Cell, 23*: 175 (1981), and other cell lines capable of expressing a compatible vector, for example, the C127, 3T3, CHO, HeLa and BHK cell lines. Mammalian expression vectors will comprise an origin of replication, a suitable promoter and enhancer, and also any necessary ribosome binding sites, polyadenylation site, splice donor and acceptor sites, transcriptional termination sequences, and 5' flanking nontranscribed sequences. DNA sequences derived from the SV40 splice, and polyadenylation sites may be used to provide the required nontranscribed genetic elements.

[0103] The enzyme can be recovered and purified from recombinant cell cultures by methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Protein refolding steps can be used, as necessary, in completing configuration of the mature protein. Finally, high performance liquid chromatography (HPLC) can be employed for final purification steps.

[0104] The enzymes of the present invention may be a naturally purified product, or a product of chemical synthetic procedures, or produced by recombinant techniques from a prokaryotic or eukaryotic host (for example, by bacterial, yeast, higher plant, insect and mammalian cells in culture). Depending upon the host employed in a recombinant production procedure, the enzymes of the present invention may be glycosylated or may be non-glycosylated. Enzymes of the invention may or may not also include an initial methionine amino acid residue.

[0105] Esterases are a group of key enzymes in the metabolism of fats and are found in all organisms from microbes to mammals. In the hydrolysis reaction, an ester group is hydrolysed to an organic acid and an alcohol.

[0106] Esterases enantiomerically differentiate dicarboxylic diesters and diacetates of diols. Using the approach disclosed in a commonly assigned, copending provisional application Serial No. 60/008,316, filed on December 7, 1995 and entitled "Combinatorial Enzyme Development," the disclosure of which is incorporated herein by reference in its entirety, one could convert the enantiospecificity of the esterase. Further, the thermostable esterases are believed to have superior stability at higher temperatures and in organic solvents. Thus, they are better suited for use in rigorous production procees which require robust catalysts.

[0107] There are a number of industrial and scientific applications for esterases, such as those of the present invention. including:

1) Esterases are useful in the dairy industry as ripening starters for cheeses, such as the Swiss-type cheeses;

2) Esterases are useful in the pulp and paper industry for lignin removal from cellulose pulps, for lignin solubilization by cleaving the ester linkages between aromatic acids and lignin and between lignin and hemicelluloses, and for disruption of cell wall structure when used in combination with xylanase and other xylan-degrading enzymes in biopulping and biobleaching of pulps;

3) Esterases are useful in the synthesis of carbohydrate derivatives, such as sugar derivatives;

4) Esterases are useful, when combined with xylanases and cellulases, in the conversion of lignocellulosic wastes to fermentable sugars for producing a variety of chemicals and fuels;

5) Esterases are useful as research reagents in studies on plant cell wall structure, particularly the nature of covalent bonds between lignin and carbohydrate polymers in the cell wall matrix;

6) Esterases are also useful as research reagents in studies on mechanisms related to disease resistance in plants and the process of organic matter decomposition; and

7) Esterases are useful in selection of plants bred for production of highly digestible animal feeds, particularly for ruminant animals.

[0108] Antibodies generated against the enzymes corresponding to a sequence of the present invention can be obtained by direct injection of the enzymes into an animal or by administering the enzymes to an animal, preferably a nonhuman. The antibody so obtained will then bind the enzymes itself. In this manner, even a sequence encoding only a fragment of the enzymes can be used to generate antibodies binding the whole native enzymes. Such antibodies can then be used to isolate the enzyme from cells expressing that enzyme.

[0109] For preparation of monoclonal antibodies, any technique which provides antibodies produced by continuous cell line cultures can be used. Examples include the hybridoma technique (Kohler and Milstein, *Nature, 256*:495-497, 1975), the trioma technique, the human B-cell hybridoma technique (Kozbor *et al., Immunology Today 4*:72, 1983), and the EBV-hybridoma technique to produce human monoclonal antibodies (Cole *et al.*, in *Monoclonal Antibodies and Cancer Therapy,* Alan R. Liss, Inc., pp. 77-96, 1985).

[0110] Techniques described for the production of single chain antibodies (U.S. Patent 4.946.778) can be adapted to produce single chain antibodies to immunogenic enzyme products of this invention. Also, transgenic mice may be used to express humanized antibodies to immunogenic enzyme products of this invention.

[0111] Antibodies generated against an enzyme of the present invention may be used in screening for similar enzymes from other organisms and samples. Such screening techniques are known in the art, for example, one such screening assay is described in Sambrook *et al.,* Molecular Cloning: A Laboratory Manual (2d Ed.), Cold Spring Harbor Laboratory, Section 12.21-12.28 (1989) which is hereby incorporated by reference in its entirety.

[0112] The present invention will be further described with reference to the following examples; however, it is to be understood that the present invention is not limited to such examples. All parts or amounts, unless otherwise specified, are by weight.

[0113] In order to facilitate understanding of the following examples certain frequently occurring methods and/or terms will be described.

[0114] "Plasmids" are designated by a lower case "p" preceded and/or followed by capital letters and/or numbers. The starting plasmids herein are either commercially available, publicly available on an unrestricted basis, or can be constructed from available plasmids in accord with published procedures. In addition, equivalent plasmids to those described are known in the art and will be apparent to the ordinarily skilled artisan.

[0115] "Digestion" of DNA refers to catalytic cleavage of the DNA with a restriction enzyme that acts only at certain sequences in the DNA. The various restriction enzymes used herein are commercially available and their reaction conditions, cofactors and other requirements were used as would be known to the ordinarily skilled artisan. For analytical purposes, typically 1 µg of plasmid or DNA fragment is used with about 2 units of enzyme in about 20 µl of buffer solution. For the purpose of isolating DNA fragments for plasmid construction, typically 5 to 50 µg of DNA are digested with 20 to 250 units of enzyme in a larger volume. Appropriate buffers and substrate amounts for particular restriction enzymes are specified by the manufacturer. Incubation times of about 1 hour at 37' C are ordinarily used, but may vary in accordance with the supplier's instructions. After digestion the reaction is electrophoresed directly on a polyacrylamide gel to isolate the desired fragment.

[0116] Size separation of the cleaved fragments is performed using 8 percent polyacrylamide gel described by Goeddel *et al., Nucleic Acids Res., 8*:4057 (1980).

[0117] "Oligonucleotides" refers to either a single stranded polydeoxynucleotide or two complementary polydeoxynucleotide strands which may be chemically synthesized. Such synthetic oligonucleotides have no 5' phosphate and thus will not ligate to another oligonucleotide without adding a phosphate with an ATP in the presence of a kinase. A synthetic oligonucleotide will ligate to a fragment that has not been dephosphorylated.

[0118] "Ligation" refers to the process of forming phosphodiester bonds between two double stranded nucleic acid fragments (Maniatis, T., *et al., Id.*, p. 146). Unless otherwise provided, ligation may be accomplished using known buffers and conditions with 10 units of T4 DNA ligase ("ligase") per 0.5 µg of approximately equimolar amounts of the DNA fragments to be ligated.

[0119] Unless otherwise stated. transformation was performed as described in Sambrook *et al.*, Molecular Cloning: A Laboratory Manual (2d Fd.), Cold Spring Harbor Press (1989).

Example 1

Bacterial Expression and Purification of Esterases

[0120] DNA encoding the enzymes of the present invention, SEQ ID NOS:33 through 42, were initially amplified from a pBluescript vector containing the DNA by the PCR technique using the primers noted herein. The amplified sequences were then inserted into the respective PQE vector listed beneath the primer sequences. and the enzyme was expressed according to the protocols set forth herein. The 5' and 3' primer sequences for the respective genes are as follows:

Staphylothermus marinus F1-12LC

5' CCGAGAATTC ATTAAAGAGG AGAAATTAAC TATGTCTTTA AACAAGCACT CT
3' CGGAAGATCT CTATCGTTTA GTGTATGATT T

vector: pQET

Pyrodictium TAG 11-17LC

5' CCGAGAATTC ATTAAAGAGG AGAAATTAAC TATGAAACTC CTTGAGCCCA CA **EcoRI**
3' CGGAAGATCT CGCCGGTACA CCATCAGCCA C **BglII**

vector: pQET

Archaeoglobus venificus SNP6-24LC

5' CCGAGAATTC ATTAAAGAGG AGAAATTAAC TATGCCATAT GTTAGGAATG GT
3' CGGAGGTACC TTAGAACTGT GCTGAAGAAA TAAATTCGTC CATTGCTCT
3' CGGAGGTACC TTAGAACTGT GCTGAAGAAA TAAATTCGTC CATTGCTCTA TTA

vector: pQET

Aquifex pyrophilus - 28LC

5' CCGAGAATTC ATTAAAGAGG AGAAATTAAC TATGAGATTG AGGAAATTTG AAG
3' CGGAGGTACC CTATTCAGAA AGTACCTCTA A

vector: pQET

M11TL - 29LC

5' CCGAGAATTC ATTAAAGAGG AGAAATTAAC TATGTTTAAT ATCAATGTCT TT
3' CGGAAGATCT TTAAGGATTT TCCCTGGGTA G

vector: pQET

Thermococcus CL-2 - 30LC

5' CCGAGAATTC ATTAAAGAGG AGAAATTAAC TATGGAGGTT TACAAGGCCA AA
3' CGGAGGTACC TTATTGAGCC GAAGAGTACG A

vector: pQET

Aquifex VF5 - 34LC

5' CCGAGAATTC ATTAAAGAGG AGAAATTAAC TATGATTGGC AATTTGAAAT TGA EcoRI
3' CGGAGGTACC TTAAAGTGCT CTCATATCCC C KpnI

vector: pQET

Teredinibacter 42L

5' CCGAGAATTC ATTAAAGAGG AGAAATTAAC TATGCCAGCT AATGACTCAC CC
3' CGGAAGATCT TCAACAGGCT CCAAATAATT TC (without His-tag)
3' CGGAAGATCT ACAGGCTCCA AATAATTTC (with His-tag)

vector: pQE12

Archaeoglobus fulgidus VC16-16MC

5' CCGAGAATTC ATTAAAGAGG AGAAATTAAC TATGCTTGAT ATGCCAATCG AC EcoRI
3' CGGAGGTACC CTAGTCGAAG ACAAGAAGAG C KpnI

vector: pQET

Sulfolabus solfataricus P1-8LC

5' CCGAGAATTC ATTAAAGAGG AGAAATTAAC TATGCCCCAG GATCCTAGAA TT EcoRI
3' CGGAGGTACC TTAAATTTTA TCATAAAATA C KpnI

vector: pQET

[0121] The restriction enzyme sites indicated correspond to the restriction enzyme sites on the bacterial expression vector indicated for the respective gene (Qiagen, Inc. Chatsworth, CA). The pQE vector encodes antibiotic resistance (Amp^r), a bacterial origin of replication (ori). an IPTG-regulatable promoter operator (P/O), a ribosome binding site (RBS), a 6-His tag and restriction enzyme sites.

[0122] The pQE vector was digested with the restriction enzymes indicated. The amplified sequences were ligated into the respective pQE vector and inserted in frame with the sequence encoding for the RBS. The ligation mixture was then used to transform the *E. coli* strain M15/pREP4 (Qiagen, Inc.) by electroporation. M15/pREP4 contains multiple copies of the plasmid pREP4, which expresses the lacl repressor and also confers kanamycin resistance (Kan^r). Transformants were identified by their ability to grow on LB plates and ampicillin/kanamycin resistant colonies were selected. Plasmid DNA was isolated and confirmed by restriction analysis. Clones containing the desired constructs were grown overnight (O/N) in liquid culture in LB media supplemented with both Amp (100 ug/ml) and Kan (25 ug/ml). The O/N culture was used to inoculate a large culture at a ratio of 1:100 to 1:250. The cells were grown to an optical density 600 (O.D.^{600}) of between 0.4 and 0.6. IPTG ("Isopropyl-B-D-thiogalacto pyranoside") was then added to a final concentration of 1 mM. IPTG induces by inactivating the lacl repressor, clearing the P/O leading to increased gene expression. Cells were grown an extra 3 to 4 hours. Cells were then harvested by centrifugation.

[0123] The primer sequences set out above may also be employed to isolate the target gene from the deposited material by hybridization techniques described above.

Example 2

Isolation of a Selected Clone from the Deposited Genomic Clones

[0124] The two oligonucleotide primers corresponding to the gene of interest are used to amplify the gene from the deposited material. A polymerase chain reaction is carried out in 25 µl of reaction mixture with 0.1 µg of the DNA of the gene of interest. The reaction mixture is 1.5-5 mM $MgCl_2$, 0.01 % (w/v) gelatin, 20 µM each of dATP, dCTP, dGTP, dTTP. 25 pmol of each primer and 1.25 Unit of Taq polymerase. Thirty cycles of PCR (denaturation at 94°C for 1 min; annealing at 55°C for 1 min; elongation at 72°C for 1 min) are performed with the Perkin-Elmer Cetus 9600 thermal cycler. The amplified product is analyzed by agarose gel electrophoresis and the DNA band with expected molecular weight is excised and purified. The PCR product is verified to be the gene of interest by subcloning and sequencing the DNA product.

Example 3

Production of the Expression Gene Bank

[0125] Colonies containing pBluescript plasmids with random inserts from the organisms M11TL, *Thermococcus* GU5L5, and *Teredinibacter* were obtained according to the method of Hay and Short, *Strategies*, *5*:16, 1992.

Example 4

Screening for Lipase/Esterase Activity

[0126] The resulting colonies were picked with sterile toothpicks and used to singly inoculate each of the wells of 96-well microtiter plates. The wells contained 250 µL of LB media with 100 µg/mL ampicillin, 80 µg/mL methicillin, and 10% v/v glycerol (LB Amp/Meth. glycerol). The cells were grown overnight at 37°C without shaking. This constituted generation of the "Source GeneBank." Each well of the Source GeneBank thus contained a stock culture of *E. coli* cells, each of which contained a pBluescript with a unique DNA insert.

[0127] The plates of the Source GeneBank were used to multiply inoculate a single plate (the "Condensed Plate") containing in each well 200 µL of LB Amp/Meth, glycerol. This step was performed using the High Density Replicating Tool (HDRT) of the Beckman Biomek with a 1 % bleach, water, isopropanol, air-dry sterilization cycle in between each inoculation. Each well of the Condensed Plate thus contained 10 to 12 different pBluescript clones from each of the source library plates. The Condensed Plate was grown for 16 hours at 37°C and then used to inoculate two white 96-well Polyfiltronics microtiter daughter plates containing in each well 250 µL of LB Amp/Meth (no glycerol). The original condensed plate was put in storage -80°C. The two condensed daughter plates were incubated at 37°C for 18 hours.

[0128] The short chain esterase '600 µM substrate stock solution' was prepared as follows: 25 mg of each of the following compounds was dissolved in the appropriate volume of DMSO to yield a 25.2 mM solution. The compounds used were 4-methylumbelliferyl proprionoate, 4-methylumbelliferyl butyrate, and 4-methylumbelliferyl heptanoate. Two hundred fifty microliters of each DMSO solution was added to ca 9 mL of 50 mM, pH 7.5 Hepes buffer which contained 0.6% of Triton X-100 and 0.6 mg per mL of dodecyl maltoside (Anatrace). The volume was taken to 10.5 mL with the above Hepes buffer to yield a slightly cloudy suspension.

[0129] The long chain '600 µM substrate stock solution' was prepared as follows: 25 mg of each of the following compounds was dissolved in DMSO to 25.2 mM as above. The compounds used were 4-methylumbelliferyl elaidate, 4-methylumbelliferyl palmitate, 4-methylumbelliferyl oleate, and 4-methylumbelliferyl stearate. All required brief warming in a 70°C bath to achieve dissolution. Two hundred fifty microliters of each DMSO solution was added to the Hepes buffer and diluted to 10.5 mL as above. All seven umbelliferones were obtained from Sigma Chemical Co.

[0130] Fifty µL of the long chain esterase or short chain esterase '600 µM substrate stock solution' was added to each of the wells of a white condensed plate using the Biomek to yield a final concentration of substrate of about 100 µM.. The fluorescence values were recorded (excitation = 326 nm, emission = 450 nm) on a plate-reading fluorometer immediately after addition of the substrate. The plate was incubated at 70°C for 60 minutes in the case of the long chain substrates, and 30 minutes at RT in the case of the short chain substrates. The fluorescence values were recorded again. The initial and final fluorescence values were compared to determine if an active clone was present.

Example 5

Isolation and Purification of the Active Clone

[0131] To isolate the individual clone which carried the activity, the Source GeneBank plates were thawed and the individual wells used to singly inoculate a new plate containing LB Amp/Meth. As above, the plate was incubated at 37°C to grow the cells, 50 µL of 600 µM substrate stock solution was added using the Biomek and the fluorescence was determined. Once the active well from the source plate was identified, cells from this active well were streaked on agar with LB/Amp/Meth and grown overnight at 37°C to obtain single colonies. Eight single colonies were picked with a sterile toothpick and used to singly inoculate the wells of a 96-well microtiter plate. The wells contained 250 µL of LB Amp/Meth. The cells were grown overnight at 37°C without shaking. A 200 µL aliquot was removed from each well and assayed with the appropriate long or short chain substrates as above. The most active clone was identified and the remaining 50 µL of culture was used to streak an agar plate with LB/Amp/Meth. Eight single colonies were picked, grown and assayed as above. The most active clone was used to inoculate 3 mL cultures of LB/Amp/Meth, which were grown overnight. The plasmid DNA was isolated from the cultures and utilized for sequencing.

SEQUENCE LISTING

[0132]

- (1) GENERAL INFORMATION:
  - (i) APPLICANTS:
    - ROBERTSON, Daniel E.
    - MURPHY, Dennis
    - REID, John
    - MAFFIA, Anthony
    - LINK, Steven
    - SWANSON, Ronald V.
    - WARREN, Patrick V.
    - KOSMOTKA, Anna
    - CALLEN, Walter
  - (ii) TITLE OF INVENTION:
    - ESTERASES
  - (iii) NUMBER OF SEQUENCES: 62
  - (iv) CORRESPONDENCE ADDRESS:
    - (A) ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI, STEWART & OLSTEIN
    - (B) STREET: 6 BECKER FARM ROAD
    - (C) CITY: ROSELAND
    - (D) STATE: NEW JERSEY
    - (E) COUNTRY: USA
    - (F) ZIP: 07068
  - (v) COMPUTER READABLE FORM:
    - (A) MEDIUM TYPE: 3.5 INCH DISKETTE
    - (B) COMPUTER: IBM PS/2
    - (C) OPERATING SYSTEM: MS-DOS
    - (D) SOFTWARE: WORD PERFECT 5.1
  - (vi) CURRENT APPLICATION DATA:
    - (A) APPLICATION NUMBER: Unassigned

(B) FILING DATE: Concurrently
(C) CLASSIFICATION: Unassigned

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: HERRON, CHARLES J.
(B) REGISTRATION NUMBER: 28,019
(C) REFERENCE/DOCKET NUMBER: 331400-39

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 201-994-1700
(B) TELEFAX: 201-994-1744

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 52 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS : SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
CCGAGAATTC ATTAAAGAGG AGAAATTAAC TATGTCTTTA AACAAGCACT CT          52
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 31 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
CGGAAGATCT CTATCGTTTA GTGTATGATT T                                 31
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 52 NUCLBOTIDES
(B) TYPE: NUCLEIC ACID

(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CCGAGAATTC ATTAAAGAGG AGAAATTAAC TATGAAACTC CTTGAGCCCA CA 52

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 31 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CGGAAGATCT CGCCGGTACA CCATCAGCCA C 31

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 52 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CCGAGAATTC ATTAAAGAGG AGAAATTAAC TATGCCATAT GTTAGGAATG GT 52

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 53 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECUM: TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CGGAGGTACC TTAGAACTGT GCTGAAGAAA TAAATTCGTC CATTGCTCTA TTA 53

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 49 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CGGAGGTACC TTAGAACTGT GCTGAAGAAA TAAATTCGTC CATTGCTCT 49

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 53 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CCGAGAATTC ATTAAAGAGG AGAAATTAAC TATGAGATTG AGGAAATTTG AAG 53

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 31 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CGGAGGTACC CTATTCAGAA AGTACCTCTA A 31

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 52 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CCGAGAATTC ATTAAAGAGG AGAAATTAAC TATGTTTAAT ATCAATGTCT TT 52

(2) INFORMATION FOR SEO ID NO:11:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 31 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

CGGAAGATCT TTAAGGATTT TCCCTGGGTA G 31

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 52 NUCLEOTIDES
(B) TYPE; NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

CCGAGAATTC ATTAAAGAGG AGAAATTAAC TATGGAGGTT TACAAGGCCA AA 52

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 31 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CGGAGGTACC TTATTGAGCC GAAGAGTACG A 31

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 53 NUCLEOTIDES

(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

```
CCGAGAATTC ATTAAAGAGG AGAAATTAAC TATGATTGGC AATTTGAAAT TGA                 53
```

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 31 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

```
CGGAGGTACC TTAAAGTGCT CTCATATCCC C                                         31
```

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 31 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

```
CCGAGAATTC ATTAAAGAGG AGAAATTAAC TATGCCAGCT AATGACTCAC CC                  52
```

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 32 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

```
CGGAAGATCT TCAACAGGCT CCAAATAATT TC                                        32
```

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 29 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CGGAAGATCT ACAGGCTCCA AATAATTTC **29**

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 52 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE : cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

CCGAGAATTC ATTAAAGAGG AGAAATTAAC TATGCTTGAT ATGCCAATCG AC **52**

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 31 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

CGGAGGTACC CTAGTCGAAC AGAAGAAGAG C **31**

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 52 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

```
CCGAGAATTC ATTAAAGAGG AGAAATTAAC TATGCCCCTA GATCCTAGAA TT            52
```

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 31 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

```
CGGAGGTACC TTAAATTTTA TCATAAAATA C                                   31
```

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 555 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: GENOMIC DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

```
ATG TCT TTA AAC AAG CAC TCT TGG ATG GAT ATG ATA ATA TTT ATT CTC      48
Met Ser Leu Asn Lys His Ser Trp Met Asp Met Ile Ile Phe Ile Leu
 1               5                  10                  15

AGC TTT TCT TTC CCA TTA ACA ATG ATC GCA TTA GCT ATC TCT ATG TCG      96
Ser Phe Ser Phe Pro Leu Thr Met Ile Ala Leu Ala Ile Ser Met Ser
            20                  25                  30

TCA TGG TTT AAT ATA TGG AAT AAT GCA TTA AGC GAT CTA GGA CAT GCT     144
Ser Trp Phe Asn Ile Trp Asn Asn Ala Leu Ser Asp Leu Gly His Ala
        35                  40                  45

GTT AAA AGC AGT GTT GCT CCA ATA TTC AAT CTA GGT CTT GCA ATT GGT     192
Val Lys Ser Ser Val Ala Pro Ile Phe Asn Leu Gly Leu Ala Ile Gly
    50                  55                  60

GGG ATA CTA ATT GTT ATA GTT GGT TTA AGA AAT CTT TAT TCG TGG AGT     240
Gly Ile Leu Ile Val Ile Val Gly Leu Arg Asn Leu Tyr Ser Trp Ser
65                  70                  75                  80

AGA GTT AAA GGA TCT TTA ATC ATA TCC ATG GGT GTA TTT CTT AAC TTA     288
Arg Val Lys Gly Ser Leu Ile Ile Ser Met Gly Val Phe Leu Asn Leu
                85                  90                  95

ATA GGG GTT TTC GAC GAA GTA TAT GGT TGG ATA CAT TTC CTA GTC TCA     336
Ile Gly Val Phe Asp Glu Val Tyr Gly Trp Ile His Phe Leu Val Ser
            100                 105                 110

GTA TTG TTT TTC TTA TCA ATA ATA GCA TAT TTC ATA GCT ATA TCA ATA     384
Val Leu Phe Phe Leu Ser Ile Ile Ala Tyr Phe Ile Ala Ile Ser Ile
        115                 120                 125

CTT GAC AAA TCA TGG ATA GCT GTT CTA CTA ATA ATA GGT CAT ATT GCA     432
Leu Asp Lys Ser Trp Ile Ala Val Leu Leu Ile Ile Gly His Ile Ala
    130                 135                 140

ATG TGG TAT CTA CAC TTT GCT TCA GAG ATT CCG AGA GGT GCG GCT ATT     480
Met Trp Tyr Leu His Phe Ala Ser Glu Ile Pro Arg Gly Ala Ala Ile
145                 150                 155                 160

CCC GAG TTA TTA GCG GTA TTC TCG TTT TTA CCA TTC TAT ATA AGA GAC     528
Pro Glu Leu Leu Ala Val Phe Ser Phe Leu Pro Phe Tyr Ile Arg Asp
                165                 170                 175

TAT TTT AAA TCA TAC ACT AAA CGA TAG                                 576
Tyr Phe Lys Ser Tyr Thr Lys Arg
            180
```

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 1041 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: GENOMIC DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

```
ATG AAA CTC CTT GAG CCC ACA AAT ACC TCC TAC ACG CTG TTA CAG GAT      48
Met Lys Leu Leu Glu Pro Thr Asn Thr Ser Tyr Thr Leu Leu Gln Asp
 1               5                  10                  15

TTA GCA TTG CAT TTT GCA TTT TAC TGG TTT CTG GCC GTG TAT ACG TGG
 96
Leu Ala Leu His Phe Ala Phe Tyr Trp Phe Leu Ala Val Tyr Thr Trp
            20                  25                  30

TTA CCC GGT GTC CTA GTC CGG GGC GTA GCT GTG GAC ACA GGG GTG GCT     144
Leu Pro Gly Val Leu Val Arg Gly Val Ala Val Asp Thr Gly Val Ala
        35                  40                  45

CGG GTG CCT GGG CTC GGC CGG CGC GGT AAG AGG CTG CTC CTG GCC GCT     192
Arg Val Pro Gly Leu Gly Arg Arg Gly Lys Arg Leu Leu Leu Ala Ala
    50                  55                  60

GTG GCT GTC TTG GCG CTT GTT GTG TCC GTT GTT GTC CCG GCT TAT GTG     240
Val Ala Val Leu Ala Leu Val Val Ser Val Val Val Pro Ala Tyr Val
65                  70                  75                  80

GCG TAT AGT AGT CTG CAC CCG GAG AGC TGT CGG CCC GTT GCG CCG GAG     288
Ala Tyr Ser Ser Leu His Pro Glu Ser Cys Arg Pro Val Ala Pro Glu
                85                  90                  95

GGG CTC ACC TAC AAA GAG TTC AGC GTG ACC GCG GAG GAT GGC TTG GTG     336
Gly Leu Thr Tyr Lys Glu Phe Ser Val Thr Ala Glu Asp Gly Leu Val
            100                 105                 110

GTT CGG GGC TGG GTG CTG GGC CCC GGC GCT GGG GGC AAC CCG GTG TTC     384
Val Arg Gly Trp Cal Leu Gly Pro Gly Ala Gly Gly Asn Pro Val Phe
        115                 120                 125

GTT TTG ATG CAC GGG TAT ACT GGG TGC CGC TCG GCG CCC TAC ATG GCT     432
Val Leu Met His Gly Tyr Thr Gly Cys Arg Ser Ala Pro Tyr Met Ala
    130                 135                 140

GTG CTG GCC CGG GAG CTC GTG GAG TGG GGG TAC CCG GTG GTT GTG TTC     480
Val Leu Ala Arg Glu Leu Val Glu Trp Gly Tyr Pro Val Val Val Phe
145                 150                 155                 160

GAC TTC CGG GGC CAC GGG GAG AGC GGG GGC TCG ACG ACG ATT GGG CCC     528
Asp Phe Arg Gly His Gly Glu Ser Gly Gly Ser Thr Thr Ile Gly Pro
```

```
                    165                 170                 175

CGG GAG GTG CTG GAT GCC CGG GCT GTG GTG GGC TAT GTC TCG GAG CGG       576
Arg Glu Val Leu Asp Ala Arg Ala Val Val Gly Tyr Val Ser Glu Arg
            180                 185                 190

TTC CCC GGC CGC CGG ATA ATA TTG GTG GGG TTC AGT ATG GGC GGC GCT       624
Phe Pro Gly Arg Arg Ile Ile Leu Val Gly Phe Ser Met Gly Gly Ala
        195                 200                 205

GTA GCG ATC GTG GAG GGT GCT GGG GAC CCG CGG GTC TAC GCG GTG GCT       672
Val Ala Ile Val Glu Gly Ala Gly Asp Pro Arg Val Tyr Ala Val Ala
    210                 215                 220

GCT GAT AGC CCG TAC TAT AGG CTC CGG GAC GTC ATA CCC CGG TGG CTG       720
Ala Asp Ser Pro Tyr Tyr Arg Leu Arg Asp Val Ile Pro Arg Trp Leu
225                 230                 235                 240

GAG TAC AAG ACG CCG CTG CCG GGC TGG GTG GGT GTG CTG GCC GGG TTC       768
Glu Tyr Lys Thr Pro Leu Pro Gly Trp Val Gly Val Leu Ala Gly Phe
                245                 250                 255

TAC GGG AGG CTG ATG GCG GGC GTT GAC CTC GGC TTC GGC CCC GCT GGG       816
Tyr Gly Arg Leu Met Ala Gly Val Asp Leu Gly Phe Gly Pro Ala Gly
            260                 265                 270

GTG GAG CGC GTG GAT AAG CCG TTG CTG GTG GTG TAT GGG CCC CGG GAC       864
Val Glu Arg Val Asp Lys Pro Leu Leu Val Val Tyr Gly Pro Arg Asp
        275                 280                 285

CCG CTG GTG ACG CGG GAC GAG GCG AGG AGC CTG GCG TCC CGT AGC CCG       912
Pro Leu Val Thr Arg Asp Glu Ala Arg Ser Leu Ala Ser Arg Ser Pro
    290                 295                 300

TGT GGC CGT CTC GTC GAG GTT CCT GGG GCT GGC CAC GTG GAG GCC GTG       960
Cys Gly Arg Leu Val Glu Val Pro Gly Ala Gly His Val Glu Ala Val
305                 310                 315                 320

GAT GTG CTC GGG CCG GGC CGC TAC GCA GAC ATG CTG ATA GAG CTG GCG      1008
Asp Val Leu Gly Pro Gly Arg Tyr Ala Asp Met Leu Ile Glu Leu Ala
                325                 330                 335

CAC GAG GAG TGC CCT CCG GGG GCC GGT GGC TGA                          1029
His Glu Glu Cys Pro Pro Gly Ala Gly Gly
            340                 345
```

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 789 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: GENOMIC DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

```
ATG CCA TAT GTT AGG AAT GGT GGT GTA AAT ATC TAT TAT GAA CTG GTG        48
Met Pro Tyr Val Arg Asn Gly Gly Val Asn Ile Tyr Tyr Glu Leu Val
 1               5                   10                  15

GAT GGA CCT GAG CCA CCA ATT GTC TTT GTT CAC GGA TGG ACA GCA AAT        96
Asp Gly Pro Glu Pro Pro Ile Val Phe Val His Gly Trp Thr Ala Asn
             20                  25                  30

ATG AAT TTT TGG AAA GAG CAA AGA CGT TAT TTT GCA GGC AGG AAT ATG       144
Met Asn Phe Trp Lys Glu Gln Arg Arg Tyr Phe Ala Gly Arg Asn Met
         35                  40                  45

ATG TTG TTT GTC GAT AAC AGA GGT CAT GGC AGG TCC GAT AAG CCA CTT       192
Met Leu Phe Val Asp Asn Arg Gly His Gly Arg Ser Asp Lys Pro Leu
     50                  55                  60

GGA TAC GAT TTC TAC AGA TTT GAG AAC TTC ATT TCA GAT TTA GAT GCG       240
Gly Tyr Asp Phe Tyr Arg Phe Glu Asn Phe Ile Ser Asp Leu Asp Ala
 65                  70                  75                  80

GTT GTT AGG GAG ACT GGA GTG GAG AAA TTT GTT CTC GTC GGA CAT TCA       288
Val Val Arg Glu Thr Gly Val Glu Lys Phe Val Leu Val Gly His Ser
                 85                  90                  95

TTC GGA ACA ATG ATC TCT ATG AAG TAC TGT TCG GAG TAT CGG AAT CGG       336
Phe Gly Thr Met Ile Ser Met Lys Tyr Cys Ser Glu Tyr Arg Asn Arg
            100                 105                 110

GTT CTT GCT CTA ATC CTC ATA GGT GGT GGG AGC AGA ATA AAG CTT CTA       384
Val Leu Ala Leu Ile Leu Ile Gly Gly Gly Ser Arg Ile Lys Leu Leu
        115                 120                 125

CAC AGA ATT GGA TAT CCT TTA GCA AAG ATT CTT GCA TCC ATT GCA TAC       432
His Arg Ile Gly Tyr Pro Leu Ala Lys Ile Leu Ala Ser Ile Ala Tyr
    130                 135                 140

AAG AAG TCT TCA AGA TTG GTC GCA GAT CTT TCC TTT GGC AAA AAT GCT       480
Lys Lys Ser Ser Arg Leu Val Ala Asp Leu Ser Phe Gly Lys Asn Ala
145                 150                 155                 160

GGT GAA CTT AAA GAG TGG GGA TGG AAA CAG GCA ATG GAT TAT ACA CCC       528
Gly Glu Leu Lys Glu Trp Gly Trp Lys Gln Ala Met Asp Tyr Thr Pro
                165                 170                 175

TCC TAC GTG GCA ATG TAC ACG TAC AGA ACT CTA ACG AAA GTG AAT CTT       576
Ser Tyr Val Ala Met Tyr Thr Tyr Arg Thr Leu Thr Lys Val Asn Leu
            180                 185                 190

GAA AAT ATC TTG GAG AAA ATA GAC TGT CCA ACA CTG ATT ATC GTT GGA       624
Glu Asn Ile Leu Glu Lys Ile Asp Cys Pro Thr Leu Ile Ile Val Gly
        195                 200                 205

GAA GAG GAT GCA CTA TTG CCC GTT AGC AAA TCA GTT GAG CTG AGC AGG       672
Glu Glu Asp Ala Leu Leu Pro Val Ser Lys Ser Val Glu Leu Ser Arg
    210                 215                 220

AGG ATA GAA AAC TCA AAG CTT GTG ATC ATC CCA AAC TCG GGG CAT TGC       720
Arg Ile Glu Asn Ser Lys Leu Val Ile Ile Pro Asn Ser Gly His Cys
225                 230                 235                 240

GTA ATG CTT GAG AGT CCA AGT GAG GTT AAT AGA GCA ATG GAC GAA TTC       768
Val Met Leu Glu Ser Pro Ser Glu Val Asn Arg Ala Met Asp Glu Phe
                245                 250                 255

ATT TCT TCA GCA CAG TTC TAA                                           774
Ile Ser Ser Ala Gln Phe
            260
```

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 756 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULAR TYPE: GENOMIC DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

```
TTG AGA TTG AGG AAA TTT GAA GAG ATA AAC CTC GTT CTT TCG GGA GGA       48
Leu Arg Leu Arg Lys Phe Glu Glu Ile Asn Leu Val Leu Ser Gly Gly
 1               5                   10                  15

GCT GCA AAG GGC ATA GCC CAC ATA GGT GTT TTG AAA GCT ATA AAC GAG       96
Ala Ala Lys Gly Ile Ala His Ile Gly Val Leu Lys Ala Ile Asn Glu
             20                  25                  30

CTC GGT ATA AGG GTG AGG GCT TTA AGC GGG GTG AGC GCC GGG GCA ATC      144
Leu Gly Ile Arg Val Arg Ala Leu Ser Gly Val Ser Ala Gly Ala Ile
         35                  40                  45

GTT TCG GTC TTT TAT GCC TCA GGC TAC TCC CCT GAA GGG ATG TTC AGC      192
Val Ser Val Phe Tyr Ala Ser Gly Tyr Ser Pro Glu Gly Met Phe Ser
     50                  55                  60

CTT CTG AAG AGG GTA AAC TGG CTG AAG CTG TTT AAG TTC AAG CCA CCT      240
Leu Leu Lys Arg Val Asn Trp Leu Lys Leu Phe Lys Phe Lye Pro Pro
 65                  70                  75                  80

CTG AAG GGA TTG ATA GGG TGG GAG AAG GCT ATA AGA TTC CTT GAG GAA      288
Leu Lys Gly Leu Ile Gly Trp Glu Lys Ala Ile Arg Phe Leu Glu Glu
                 85                  90                  95

GTT CTC CCT TAC AGG AGA ATA GAA AAA CTT GAG ATA CCG ACG TAT ATA      336
Val Leu Pro Tyr Arg Arg Ile Glu Lys Leu Glu Ile Pro Thr Tyr Ile
            100                 105                 110

TGC GCG ACG GAT TTA TAC TCG GGA AGG GCT CTA TAC CTC TCG GAA GGG      384
Cys Ala Thr Asp Leu Tyr Ser Gly Arg Ala Leu Tyr Leu Ser Glu Gly
        115                 120                 125

AGT TTA ATC CCC GCA CTT CTC GGC AGC TGT GCA ATT CCC GGC ATA TTT      432
Ser Leu Ile Pro Ala Leu Leu Gly Ser Cys Ala Ile Pro Gly Ile Phe
    130                 135                 140

GAA CCC GTT GAG TAT AAG AAT TAC TTG CTC GTT GAC GGA GGT ATA GTT      480
Glu Pro Val Glu Tyr Lys Asn Tyr Leu Leu Val Asp Gly Gly Ile Val
145                 150                 155                 160

AAC AAC CTT CCC GTT GAG CCC TTT CAG GAA AGC GGT ATT CCC ACC GTT      528
Asn Asn Leu Pro Val Glu Pro Phe Gln Glu Ser Gly Ile Pro Thr Val
                165                 170                 175

TGC GTT GAT GTC CTT CCC ATA GAG CCG GAA AAG GAT ATA AAG AAC ATT      576
Cys Val Asp Val Leu Pro Ile Glu Pro Glu Lys Asp Ile Lys Asn Ile
            180                 185                 190
```

```
CTT CAC ATC CTT TTG AGG AGC TTC TTT CTT GCG GTC CGC TCA AAC TCC    624
Leu His Ile Leu Leu Arg Ser Phe Phe Leu Ala Val Arg Ser Asn Ser
        195                 200                 205

GAA AAG AGA AAG GAG TTT TGT GAC CTC GTT ATA GTT CCT GAG CTT GAG    672
Glu Lys Arg Lys Glu Phe Cys Asp Leu Val Ile Val Pro Glu Leu Glu
    210                 215                 220

GAG TTC ACA CCC CTT GAT GTT AGA AAA GCG GAC CAA ATA ATG GAG AGG    720
Glu Phe Thr Pro Leu Asp Val Arg Lys Ala Asp Gln Ile Met Glu Arg
225                 230                 235                 240

GGA TAC ATA AAG GCC TTA GAG GTA CTT TCT GAA TAG                    768
Gly Tyr Ile Lys Ala Leu Glu Val Leu Ser Glu
                245                 250
```

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 894 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULAR TYPE: GENOMIC DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

```
ATG TTT AAT ATC AAT GTC TTT GTT AAT ATA TCT TGG CTG TAT TTT TCA    48
Met Phe Asn Ile Asn Val Phe Val Asn Ile Ser Trp Leu Tyr Phe Ser
1               5                   10                  15

GGG ATA GTT ATG AAG ACT GTG GAA GAG TAT GCG CTA CTT GAA ACA GGC    96
Gly Ile Val Met Lys Thr Val Glu Glu Tyr Ala Leu Leu Glu Thr Gly
            20                  25                  30

GTA AGA GTG TTT TAT CGG TGT GTA ATC CCG GAG AAA GCT TTT AAC ACT    144
Val Arg Val Phe Tyr Arg Cys Val Ile Pro Glu Lys Ala Phe Asn Thr
        35                  40                  45

TTG ATA ATA GGT TCA CAC GGA TTG GGG GCG CAC AGT GGA ATC TAC ATT    192
Leu Ile Ile Gly Ser His Gly Leu Gly Ala His Ser Gly Ile Tyr Ile
    50                  55                  60

AGT GTT GCT GAA GAA TTT GCT AGG CAC GGA TTT GGA TTC TGC ATG CAC    240
Ser Val Ala Glu Glu Phe Ala Arg His Gly Phe Gly Phe Cys Met His
65                  70                  75                  80

GAT CAA AGG GGA CAT GGG AGA ACG GCA AGC GAT AGA GAA AGA GGG TAT    288
Asp Gln Arg Gly His Gly Arg Thr Ala Ser Asp Arg Glu Arg Gly Tyr
                85                  90                  95

GTG GAG GGC TTT CAC AAC TTC ATA GAG GAT ATG AAG GCC TTC TCC GAT    336
Val Glu Gly Phe His Asn Phe Ile Glu Asp Met Lys Ala Phe Ser Asp
            100                 105                 110

TAT GCC AAG TGG CGC GTG GGA GGT GAC GAA ATA ATA TTG CTA GGA CAC    384
Tyr Ala Lys Trp Arg Val Gly Gly Asp Glu Ile Ile Leu Leu Gly His
        115                 120                 125

AGT ATG GGC GGG CTG ATA GCG CTC TTA ACA GTT GCA ACT TAT AAA GAA    432
Ser Met Gly Gly Leu Ile Ala Leu Leu Thr Val Ala Thr Tyr Lys Glu
    130                 135                 140

ATC GCC AAG GGA GTT ATC GCG CTA GCC CCG GCC CTC CAA ATC CCC TTA    480
```

```
Ile Ala Lys Gly Val Ile Ala Leu Ala Pro Ala Leu Gln Ile Pro Leu
145                 150                 155                 160

ACC CCG GCT AGA AGA CTT GTT CTA AGC CTC GCG TCA AGG CTT GCC CCG       528
Thr Pro Ala Arg Arg Leu Val Leu Ser Leu Ala Ser Arg Leu Ala Pro
                165                 170                 175

CAT TCT AAG ATC ACC TTA CAA AGG AGA TTG CCG CAG AAA CCA GAG GGT       576
His Ser Lys Ile Thr Leu Gln Arg Arg Leu Pro Gln Lys Pro Glu Gly
            180                 185                 190

TTT CAA AGA GCA AAA GAT ATA GAA TAC AGT CTG AGT GAA ATA TCA GTC       624
Phe Gln Arg Ala Lys Asp Ile Glu Tyr Ser Leu Ser Glu Ile Ser Val
        195                 200                 205

AAG CTC GTG GAC GAA ATG ATT AAA GCA TCA TCT ATG TTC TGG ACC ATA       672
Lys Leu Val Asp Glu Met Ile Lys Ala Ser Ser Met Phe Trp Thr Ile
    210                 215                 220

GCA GGG GAA ATT AAT ACT CCC GTC CTG CTT ATT CAT GGG GAA AAA GAC       720
Ala Gly Glu Ile Asn Thr Pro Val Leu Leu Ile His Gly Glu Lys Asp
225                 230                 235                 240

AAT GTC ATA CCT CCG GAG GCG AGC AAA AAA GCC TAC CAA TTA ATA CCT       768
Asn Val Ile Pro Pro Glu Ala Ser Lys Lys Ala Tyr Gln Leu Ile Pro
                245                 250                 255

TCA TTC CCT AAA GAG TTG AAA ATA TAC CCC GAT CTT GGA CAC AAC TTG       816
Ser Phe Pro Lys Glu Leu Lys Ile Tyr Pro Asp Leu Gly His Asn Leu
            260                 265                 270

TTT TTT GAA CCA GGC GCG GTG AAA ATC GTC ACA GAC ATT GTA GAG TGG       864
Phe Phe Glu Pro Gly Ala Val Lys Ile Val Thr Asp Ile Val Glu Trp
        275                 280                 285

GTT AAG AAT CTA CCC AGG GAA AAT CCT TAA                               874
Val Lys Asn Leu Pro Arg Glu Asn Pro
    290                 295
```

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 789 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: GENOMIC DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

```
ATG GAG GTT TAC AAG GCC AAA TTC GGC GAA GCA AAG CTC GGC TGG GTC        48
Met Glu Val Tyr Lys Ala Lys Phe Gly Glu Ala Lys Leu Gly Trp Val
1               5                   10                  15

GTT CTG GTT CAT GGC CTC GGC GAG CAC AGC GGA AGG TAT GGA AGA CTG        96
Val Leu Val His Gly Leu Gly Glu His Ser Gly Arg Tyr Gly Arg Leu
            20                  25                  30

ATT AAG GAA CTC AAC TAT GCC GGC TTT GGA GTT TAC ACC TTC GAC TGG       144
Ile Lys Glu Leu Asn Tyr Ala Gly Phe Gly Val Tyr Thr Phe Asp Trp
        35                  40                  45

CCC GGC CAC GGG AAG AGC CCG GGC AAG AGA GGG CAC ACG AGC GTC GAG       192
```

```
Pro Gly His Gly Lys Ser Pro Gly Lys Arg Gly His Thr Ser Val Glu
    50                  55                  60

GAG GCG ATG GAA ATC ATC GAC TCG ATA ATC GAG GAG ATC AGG GAG AAG      240
Glu Ala Met Glu Ile Ile Asp Ser Ile Ile Glu Glu Ile Arg Glu Lys
 65                  70                  75                  80

CCC TTC CTC TTC GGC CAC AGC CTC GGT GGT CTA ACT GTC ATC AGG TAC      288
Pro Phe Leu Phe Gly His Ser Leu Gly Gly Leu Thr Val Ile Arg Tyr
                 85                  90                  95

GCT GAG ACG CGG CCC GAT AAA ATA CGG GGA TTA ATA GCT TCC TCG CCT      336
Ala Glu Thr Arg Pro Asp Lys Ile Arg Gly Leu Ile Ala Ser Ser Pro
            100                 105                 110

GCC CTC GCC AAG AGC CCG GAA ACG CCG GGC TTC ATG GTG GCC CTC GCG      384
Ala Leu Ala Lys Ser Pro Glu Thr Pro Gly Phe Met Val Ala Leu Ala
        115                 120                 125

AAG TTC CTT GGA AAG ATC GCC CCG GGA GTT GTT CTC TCC AAC GGC ATA      432
Lys Phe Leu Gly Lys Ile Ala Pro Gly Val Val Leu Ser Asn Gly Ile
    130                 135                 140

AAG CCG GAA CTC CTC TCG AGG AAC AGG GAC GCC GTG AGG AGG TAC GTT      480
Lys Pro Glu Leu Leu Ser Arg Asn Arg Asp Ala Val Arg Arg Tyr Val
145                 150                 155                 160

GAA GAC CCA CTC GTC CAC GAC AGG ATT TCG GCC AAG CTG GGA AGG AGC      528
Glu Asp Pro Leu Val His Asp Arg Ile Ser Ala Lys Leu Gly Arg Ser
                165                 170                 175

ATC TTC GTG AAC ATG GAG CTG GCC CAC AGG GAG GCG GAC AAG ATA AAA      576
Ile Phe Val Asn Met Glu Leu Ala His Arg Glu Ala Asp Lys Ile Lys
            180                 185                 190

GTC CCG ATC CTC CTT CTG ATC GGC ACT GGC GAT GTA ATA ACC CCG CCT      624
Val Pro Ile Leu Leu Leu Ile Gly Thr Gly Asp Val Ile Thr Pro Pro
        195                 200                 205

GAA GGC TCA CGC AGA CTC TTC GAG GAG CTG GCC GTC GAG AAC AAA ACC      672
Glu Gly Ser Arg Arg Leu Phe Glu Glu Leu Ala Val Glu Asn Lys Thr
    210                 215                 220

CTG AGG GAG TTC GAG GGG GCG TAC CAC GAG ATA TTT GAA GAC CCC GAG      720
Leu Arg Glu Phe Glu Gly Ala Tyr His Glu Ile Phe Glu Asp Pro Glu
225                 230                 235                 240

TGG GCC GAG GAG TTC CAC GAA ACA ATT GTT AAG TGG CTG GTT GAA AAA      768
Trp Ala Glu Glu Phe His Glu Thr Ile Val Lys Trp Leu Val Glu Lys
                245                 250                 255

TCG TAC TCT TCG GCT CAA TAA                                          775
Ser Tyr Ser Ser Ala Gln
            260
```

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 750 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: GENOMIC DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

```
TTG ATT GGC AAT TTG AAA TTG AAG AGG TTT GAA GAG GTT AAC TTA GTT       48
Leu Ile Gly Asn Leu Lys Leu Lys Arg Phe Glu Glu Val Asn Leu Val
1               5                   10                  15

CTT TCG GGA GGG GCT GCC AAG GGT ATC GCC CAT ATA GGT GTT TTA AAA       96
Leu Ser Gly Gly Ala Ala Lys Gly Ile Ala His Ile Gly Val Leu Lys
            20                  25                  30

GCT CTG GAA GAG CTC GGT ATA AAG GTA AAG AGG CTC AGC GGG GTA AGT       144
Ala Leu Glu Glu Leu Gly Ile Lys Val Lys Arg Leu Ser Gly Val Ser
        35                  40                  45

GCT GGA GCT ATC GTT TCC GTC TTT TAC GCT TCG GGC TAC ACT CCC GAC       192
Ala Gly Ala Ile Val Ser Val Phe Tyr Ala Ser Gly Tyr Thr Pro Asp
    50                  55                  60

GAG ATG TTA AAA CTC CTG AAA GAG GTA AAC TGG CTC AAA CTT TTT AAG       240
Glu Met Leu Lys Leu Leu Lys Glu Val Asn Trp Leu Lys Leu Phe Lys
65                  70                  75                  80

TTC AAA ACA CCG AAA ATG GGC TTA ATG GGG TGG GAG AAG GCT GCA GAG       288
Phe Lys Thr Pro Lys Met Gly Leu Met Gly Trp Glu Lys Ala Ala Glu
                85                  90                  95

TTT TTG GAA AAA GAG CTC GGA GTT AAG AGG CTG GAA GAC CTG AAC ATA       336
Phe Leu Glu Lys Glu Leu Gly Val Lys Arg Leu Glu Asp Leu Asn Ile
            100                 105                 110

CCA ACC TAT CTT TGC TCG GCG GAT CTG TAC ACG GGA AAG GCT CTT TAC       384
Pro Thr Tyr Leu Cys Ser Ala Asp Leu Tyr Thr Gly Lys Ala Leu Tyr
        115                 120                 125

TTC GGC AGA GGT GAC TTA ATT CCC GTG CTT CTC GGA AGT TGT TCC ATA       432
Phe Gly Arg Gly Asp Leu Ile Pro Val Leu Leu Gly Ser Lys Ser Ile
    130                 135                 140

CCC GGG ATT TTT GAA CCA GTT GAG TAC GAG AAT TTT CTA CTT GTT GAC       480
Pro Gly Ile Phe Glu Pro Val Glu Tyr Glu Asn Phe Leu Leu Val Asp
145                 150                 155                 160

GGA GGT ATA GTG AAC AAC CTG CCC GTA GAA CCT TTG GAA AAG TTC AAA       528
Gly Gly Ile Val Asn Asn Leu Pro Val Glu Pro Leu Glu Lys Phe Lys
                165                 170                 175

GAA CCC ATA ATC GGG GTA GAT GTG CTT CCC ATA ACT CAA GAA AGA AAG       576
Glu Pro Ile Ile Gly Val Asp Val Leu Pro Ile Thr Gln Glu Arg Lys
            180                 185                 190

ATT AAA AAT ATA CTC CAC ATC CTT ATA AGG AGC TTC TTT CTG GCG GTT       624
Ile Lys Asn Ile Leu His Ile Leu Ile Arg Ser Phe Phe Leu Ala Val
        195                 200                 205

CGT TCC AAT TCG GAA AAG AGA AAG GAG TTC TGC AAC GTA GTT ATA GAA       672
Arg Ser Asn Ser Glu Lys Arg Lys Glu Phe Cys Asn Val Val Ile Glu
    210                 215                 220

CCT CCC CTT GAA GAG TTC TCT CCT CTG GAC GTA AAT AAG GCG GAC GAG       720
Pro Pro Leu Glu Glu Phe Ser Pro Leu Asp Val Asn Lys Ala Asp Glu
225                 230                 235                 240


ATA TTC TGC GGG GAT ATG AGA GCA CTT TAA                               730
Ile Phe Cys Gly Asp Met Arg Ala Leu
                245
```

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 1017 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: GENOMIC DNA

(xi) SEQUENCE DESCRIPTION: SEQ In NO:30:

```
ATG CCA GCT AAT GAC TCA CCC ACG ATC GAC TTT AAT CCT CGC GGC ATT      48
Met Pro Ala Asn Asp Ser Pro Thr Ile Asp Phe Asn Pro Arg Gly Ile
 1               5                  10                  15

CTT CGC AAC GCT CAC GCA CAG GTT ATT TTA GCG ACT TCC GGC TTG CGC      96
Leu Arg Asn Ala His Ala Gln Val Ile Leu Ala Thr Ser Gly Leu Arg
            20                  25                  30

AAA GCG TTT TTG AAA CGC ACG CAC AAG AGC TAC CTC AGC ACT GCC CAA     144
Lys Ala Phe Leu Lys Arg Thr His Lys Ser Tyr Leu Ser Thr Ala Gln
        35                  40                  45

TGG CTG GAG CTC GAT GCC GGC AAC GGA GTT ACC TTG GCC GGA GAG CTT     192
Trp Leu Glu Leu Asp Ala Gly Asn Gly Val Thr Leu Ala Gly Glu Leu
    50                  55                  60

AAC ACA GCG CCT GCA ACT GCA TCC TCC TCC CAC CCG GCG CAC AAG AAC 240
Asn Thr Ala Pro Ala Thr Ala Ser Ser Ser His Pro Ala His Lys Asn
65                  70                  75                  80

ACT CTG GTT ATT GTG CTG CAC GGC TGG GAA GGC TCC AGC CAG TCG GCC     288
Thr Leu Val Ile Val Leu His Gly Trp Glu Gly Ser Ser Gln Ser Ala
                85                  90                  95

TAT GCG ACC TCC GCT GGC AGC ACG CTT TTC GAC AAT GGG TTC GAC ACT     336
Tyr Ala Thr Ser Ala Gly Ser Thr Leu Phe Asp Asn Gly Phe Asp Thr
            100                 105                 110

TTT CGC CTT AAT TTT CGC GAT CAC GGC GAC ACC TAC CAC TTA AAC CGC     384
Phe Arg Leu Asn Phe Arg Asp His Gly Asp Thr Tyr His Leu Asn Arg
        115                 120                 125

GGC ATA TTT AAC TCA TCG CTG ATT GAC GAA GTA GTG GGC GCA GTC AAA     432
Gly Ile Phe Asn Ser Ser Leu Ile Asp Glu Val Val Gly Ala Val Lys
    130                 135                 140

GCC ATC CAG CAG CAA ACC GAC TAC GAC AAG TAT TGC CTG ATG GGG TTC     480
Ala Ile Gln Gln Gln Thr Asp Tyr Asp Lys Tyr Cys Leu Met Gly Phe
145                 150                 155                 160

TCA CTG GGT GGG AAC TTT GCC TTG CGC GTC GCG GTG CGG GAA CAG CAT     528
Ser Leu Gly Gly Asn Phe Ala Leu Arg Val Ala Val Arg Glu Gln His
                165                 170                 175

CTC GCT AAA CCG CTA GCG GGC GTG CTC GCC GTA TGC CCG GTA CTC GAC     576
Leu Ala Lys Pro Leu Ala Gly Val Leu Ala Val Cys Pro Val Leu Asp
            180                 185                 190

CCC GCA CAC ACC ATG ATG GCC CTA AAC CGA GGT GCG TTT TTC TAC GGC     624
```

```
Pro Ala His Thr Met Met Ala Leu Asn Arg Gly Ala Phe Phe Tyr Gly
    195                 200                 205

CGC TAT TTT GCG CAT AAA TGG AAG CGC TCG TTA ACC GCA AAA CTT GCA        672
Arg Tyr Phe Ala His Lys Trp Lys Arg Ser Leu Thr Ala Lys Leu Ala
210                 215                 220                 225

GCT TTC CCA GAC TAC AAA TAC GGC AAA GAT TTA AAA TCG ATA CAC ACG        720
Ala Phe Pro Asp Tyr Lys Tyr Gly Lys Asp Leu Lys Ser Ile His Thr
                230                 235                 240

CTT GAT GAG TTA AAC AAC TAT TTC ATT CCC CGC TAC ACC GGC TTC AAC        768
Leu Asp Glu Leu Asn Asn Tyr Phe Ile Pro Arg Tyr Thr Gly Phe Asn
            245                 250                 255

TCA GTC TCC GAA TAC TTC AAA AGT TAC ACG CTC ACC GGG CAG AAG CTC        816
Ser Val Ser Glu Tyr Phe Lys Ser Tyr Thr Leu Thr Gly Gln Lys Leu
        260                 265                 270

GCG TTT CTC AAC TGC CCC AGT TAC ATT CTG GCA GCT GGC GAC GAC CCA        864
Ala Phe Leu Asn Cys Pro Ser Tyr Ile Leu Ala Ala Gly Asp Asp Pro
    275                 280                 285

ATA ATT CCA GCA TCC GAC TTT CAG AAA ATA GCC AAG CCT GCG AAT CTG        912
Ile Ile Pro Ala Ser Asp Phe Gln Lys Ile Ala Lys Pro Ala Asn Leu
290                 295                 300                 305

CAC ATA ACA GTA ACG CAA CAA GGT TCT CAT TGC GCA TAC CTG GAA AAC        960
His Ile Thr Val Thr Gln Gln Gly Ser His Cys Ala Tyr Leu Glu Asn
                310                 315                 320

CTG CAT AAA CCT AGT GCT GCC GAC AAA TAT GCG GTG AAA TTA TTT GGA      1,008
Leu His Lys Pro Ser Ala Ala Asp Lys Tyr Ala Val Lys Leu Phe Gly
            325                 330                 335

GCC TGT TGA                                                          1,111
Ala Cys
```

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 936 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: GENOMIC DNA .

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

```
ATG CTT GAT ATG CCA ATC GAC CCT GTT TAC TAC CAG CTT GCT GAG TAT      48
Met Leu Asp Met Pro Ile Asp Pro Val Tyr Tyr Gln Leu Ala Glu Tyr
 1               5                  10                  15

TTC GAC AGT CTG CCG AAG TTC GAC CAG TTT TCC TCG GCC AGA GAG TAC      96
Phe Asp Ser Leu Pro Lys Phe Asp Gln Phe Ser Ser Ala Arg Glu Tyr
            20                  25                  30

AGG GAG GCG ATA AAT CGA ATA TAC GAG GAG AGA AAC CGG CAG CTG AGC     144
Arg Glu Ala Ile Asn Arg Ile Tyr Glu Glu Arg Asn Arg Gln Leu Ser
        35                  40                  45

CAG CAT GAG AGG GTT GAA AGA GTT GAG GAC AGG ACG ATT AAG GGG AGG     192
Gln His Glu Arg Val Glu Arg Val Glu Asp Arg Thr Ile Lys Gly Arg
    50                  55                  60
```

```
AAC GGA GAC ATC AGA GTC AGA GTT TAC CAG CAG AAG CCC GAT TCC CCG       240
Asn Gly Asp Ile Arg Val Arg Val Tyr Gln Gln Lys Pro Asp Ser Pro
65                  70                  75                  80

GGT CTG GTT TAC TAT CAC GGT GGT GGA TTT GTG ATT TGC AGC ATC GAG       288
Val Leu Val Tyr Tyr His Gly Gly Gly Phe Val Ile Cys Ser Ile Glu
                85                  90                  95

TCG CAC GAC GCC TTA TGC AGG AGA AYY GCG AGA CTT TCA AAC TCT ACC       336
Ser His Asp Ala Leu Cys Arg Arg Ile Ala Arg Leu Ser Asn Ser Thr
            100                 105                 110

GTA GTC TCC GTG GAT TAC AGG CTC GCT CCT GAG CAC AAG TTT CCC CCC       384
Val Val Ser Val Asp Tyr Arg Leu Ala Pro Glu His Lys Phe Pro Ala
        115                 120                 125

CCA GTT TAT CAT TGC TAC GAT GCG ACC AAG TGG GTT GCT GAG AAC CGG       432
Ala Val Tyr Asp Cys Tyr Asp Ala Thr Lys Trp Val Ala Glu Asn Ala
    130                 135                 140

GAG GAG CTG AGG ATT GAC CCG TCA AAA ATC TTC GTT GGG GGG GAC AGT       480
Glu Glu Leu Arg Ile Asp Pro Ser Lys Ile Phe Val Gly Gly Asp Ser
145                 150                 155                 160

GCG GGA CGG AAT CTT GCC CCG GCG CTT TCA ATA ATG GCG AGA GAC AGC       528
Ala Gly Gly Asn Leu Ala Ala Ala Val Ser Ile Met Ala Arg Asp Ser
                165                 170                 175

GGA GAA GAT TTC ATA AAG CAT CAA ATT CTA ACT TAC CCC GTT GTG AAC       576
Gly Glu Asp Phe Ile Lys His Gln Ile Leu Ile Tyr Pro Val Val Asn
            180                 185                 190

TTT GTA GCC CCC ACA CCA TCG CTT CTG GAG TTT GGA GAG GGG CTG TGG       624
Phe Val Ala Pro Thr Pro Ser Leu Leu Glu Phe Gly Glu Gly Leu Trp
        195                 200                 205

ATT CTC GAC CAG AAG ATA ATG AGT TGG TTC TCG GAG CAG TAC TTC TCC       672
Ile Leu Asp Gln Lys Ile Met Ser Trp Phe Ser Glu Gln Tyr Phe Ser
    210                 215                 230

AGA GAG GAA GAT AAG TTC AAG CCC CTC GCC TCC GTA ATC TTT GCG GAC       720
Arg Glu Glu Aso Lys Phe Asn Pro Leu Ala Ser Val Ile Phe Ala Asp
235                 240                 245                 250

CTT GAG AAC CTA CCT CCT GCG CTG ATC ATA ACC GCC GAA TAC GAC CCG       768
Leu Glu Asn Leu Pro Pro Ala Leu Ile Ile Thr Ala Glu Tyr Asp Pro
                255                 260                 265

CTG AGA GAT GAA GGA GAA GTT TTC GGG CAG ATG CTG AGA AGA GCC GGT       816
Leu Arg Asp Glu Gly Glu Val Phe Gly Gln Met Leu Arg Arg Ala Gly
            270                 275                 280

GTT GAG GCG AGC ATC GTC AGA TAC AGA GGC GTG CTT CAC GGA TTC ATC       864
Val Glu Ala Ser Ile Val Arg Tyr Arg Gly Val Leu His Gly Phe Ile
        285                 290                 295

AAT TAC TAT CCC GTG CTG AAG GCT GCG AGG GAT GCG ATA AAC CAG ATT       912
Asn Tyr Tyr Pro Val Leu Lys Ala Ala Arg Asp Ala Ile Asn Gln Ile
    300                 305                 310

GCC GCT CTT CTT GTG TTC GAC TAG                                       936
Ala Ala Leu Leu Val Phe Asp
315                 320
```

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 918 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID

(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: GENOMIC DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

```
ATG CCC CTA GAT CCT AGA ATT AAA AAG TTA CTA GAA TCA GCT CTT ACT      48
Met Pro Leu Asp Pro Arg Ile Lys Lys Leu Leu Glu Ser Ala Leu Thr
                5                   10                  15

ATA CCA ATT GGT AAA GCC CCA GTA GAA GAG GTA AGA AAG ATA TTT AGG      96
Ile Pro Ile Gly Lys Ala Pro Val Glu Glu Val Arg Lys Ile Phe Arg
            20                  25                  30

CAA TTA GCG TCG GCA GCT CCC AAA GTC GAA GTT GGA AAA GTA GAA GAT     144
Gln Leu Ala Ser Ala Ala Pro Lys Val Glu Val Gly Lys Val Glu Asp
        35                  40                  45

ATA AAA ATA CCA GGC AGT GAA ACC GTT ATA AAC GCT AGA GTG TAT TTT     192
Ile Lys Ile Pro Gly Ser Glu Thr Val Ile Asn Ala Arg Val Tyr Phe
    50                  55                  60

CCG AAG AGT AGC GGT CCT TAT GGT GTT CTA GTG TAT CTT CAT GGA GGC     240
Pro Lys Ser Ser Gly Pro Tyr Gly Val Leu Val Tyr Leu His Gly Gly
65                  70                  75                  80

GGT TTT GTA ATA GGC GAT GTG GAA TCT TAT GAC CCA TTA TGT AGA GCA     288
Gly Phe Val Ile Gly Asp Val Glu Ser Tyr Asp Pro Leu Cys Arg Ala
                85                  90                  95

ATT ACA AAT GCG TGC AAT TGC GTT GTA GTA TCA GTG GAC TAT AGG TTA     336
Ile Thr Asn Ala Cys Asn Cys Val Val Val Ser Val Asp Tyr Arg Leu
            100                 105                 110

GCT CCA GAA TAC AAG TTT CCT TCT GCA GTT ATC GAT TCA TTT GAC GCT     384
Ala Pro Glu Tyr Lys Phe Pro Ser Ala Val Ile Asp Ser Phe Asp Ala
        115                 120                 125

ACT AAT TGG GTT TAT AAC AAT TTA GAT AAA TTT GAT GGA AAG ATG GGA     432
Thr Asn Trp Val Tyr Asn Asn Leu Asp Lys Phe Asp Gly Lys Met Gly
    130                 135                 140

GTT GCG ATT GCG GGA GAT AGT GCT GGA GGA AAT TTG GCA GCG GTT GTA     480
Val Ala Ile Ala Gly Asp Ser Ale Gly Gly Asn Leu Ala Ala Val Val
145                 150                 155                 160

GCT CTT CTT TCA AAG GGT AAA ATT AAT TTG AAG TAT CAA ATA CTG GTT     528
Ala Leu Leu Ser Lys Gly Lys Ile Asn Leu Lys Tyr Gln Ile Leu Val
                165                 170                 175

TAC CCA GCG GTA AGT TTA GAT AAC GTT TCA AGA TCC ATG ATA GAG TAC     576
Tyr Pro Ala Val Ser Leu Asp Asn Val Ser Arg Ser Met Ile Glu Tyr
            180                 185                 190

TCT GAT GGG TTC TTC CTT ACC AGA GAG CAT ATA GAG TGG TTC GGT TCT     624
Ser Asp Gly Phe Phe Leu Thr Arg Glu His Ile Glu Trp Phe Gly Ser
        195                 200                 205
```

```
CAA TAC TTA CGA AGC CCT GCA GAT TTG CTA GAC TTT AGG TTC TCT CCA      672
Gln Tyr Leu Arg Ser Pro Ala Asp Leu Leu Asp Phe Arg Phe Ser Pro
    210                 215                 220

ATT CTG GCG CAA GAT TTC AAC GGA TTA CCT CCA GCC TTG ATA ATA ACA      720
Ile Leu Ala Gln Asp Phe Asn Gly Leu Pro Pro Ala Leu Ile Ile Thr
225                 230                 235                 240

GCA GAA TAC GAT CCA CTA AGG GAT CAA GGA GAA GCG TAT GCA AAT AAA      768
Ala Glu Tyr Asp Pro Leu Arg Asp Gln Gly Glu Ala Tyr Ala Asn Lys
                245                 250                 255

CTA CTA CAA GCT GGA GTC TCA GTT ACT AGT GTG AGA TTT AAC AAC GTT      816
Leu Leu Gln Ala Gly Val Ser Val Thr Ser Val Arg Phe Asn Asn Val
            260                 265                 270

ATA CAC GGA TTC CTC TCA TTC TTT CCG TTG ATG GAG CAA GGA AGA GAT      864
Ile His Gly Phe Leu Ser Phe Phe Pro Leu Met Glu Gln Gly Arg Asp
        275                 280                 285

GCT ATA GGT CTG ATA GGG TCT GTG TTA AGA CGA GTA TTT TAT GAT AAA      912
Ala Ile Gly Leu Ile Gly Ser Val Leu Arg Arg Val Phe Tyr Asp Lys
    290                 295                 300

ATT TAA                                                              918
Ile
305
```

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 184 AMINO ACIDS
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

```
Met Ser Leu Asn Lys His Ser Trp Met Asp Met Ile Ile Phe Ile Leu
1               5                   10                  15

Ser Phe Ser Phe Pro Leu Thr Met Ile Ala Leu Ala Ile Ser Met Ser
            20                  25                  30

Ser Trp Phe Asn Ile Trp Asn Asn Ala Leu Ser Asp Leu Gly His Ala
        35                  40                  45

Val Lys Ser Ser Val Ala Pro Ile Phe Asn Leu Gly Leu Ala Ile Gly
    50                  55                  60

Gly Ile Leu Ile Val Ile Val Gly Leu Arg Asn Leu Tyr Ser Trp Ser
65                  70                  75                  80

Arg Val Lys Gly Ser Leu Ile Ile Ser Met Gly Val Phe Leu Asn Leu
                85                  90                  95

Ile Gly Val Phe Asp Glu Val Tyr Gly Trp Ile His Phe Leu Val Ser
            100                 105                 110

Val Leu Phe Phe Leu Ser Ile Ile Ala Tyr Phe Ile Ala Ile Ser Ile
        115                 120                 125

Leu Asp Lys Ser Trp Ile Ala Val Leu Leu Ile Ile Gly His Ile Ala
    130                 135                 140


Met Trp Tyr Leu His Phe Ala Ser Glu Ile Pro Arg Gly Ala Ala Ile
145                 150                 155                 160

Pro Glu Leu Leu Ala Val Phe Ser Phe Leu Pro Phe Tyr Ile Arg Asp
                165                 170                 175

Tyr Phe Lys Ser Tyr Thr Lys Arg
            180
```

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 346 AMINO ACIDS
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

```
Met Lys Leu Leu Glu Pro Thr Asn Thr Ser Tyr Thr Leu Leu Gln Asp
1               5                   10                  15

Leu Ala Leu His Phe Ala Phe Tyr Trp Phe Leu Ala Val TYr Thr Trp
            20                  25                  30

Leu Pro Gly Val Leu Val Arg Gly Val Ala Val Asp Thr Gly Val Ala
        35                  40                  45

Arg Val Pro Gly Leu Gly Arg Arg Gly Lys Arg Leu Leu Leu Ala Ala
    50                  55                  60

Val Ala Val Leu Ala Leu Val Val Ser Val Val Val Pro Ala Tyr Val
65                  70                  75                  80

Ala Tyr Ser Ser Leu His Pro Glu Ser Cys Arg Pro Val Ala Pro Glu
                85                  90                  95

Gly Leu Thr Tyr Lys Glu Phe Ser Val Thr Ala Glu Asp Gly Leu Val
            100                 105                 110

Val Arg Gly Trp Cal Leu Gly Pro Gly Ala Gly Gly Asn Pro Val Phe
        115                 120                 125

Val Leu Met His Gly Tyr Thr Gly Cys Arg Ser Ala Pro Tyr Met Ala
    130                 135                 140

Val Leu Ala Arg Glu Leu Val Glu Trp Gly Tyr Pro Val Val Val Phe
145                 150                 155                 160

Asp Phe Arg Gly His Gly Glu Ser Gly Gly Ser Thr Thr Ile Gly Pro
                165                 170                 175

Arg Glu Val Leu Asp Ala Arg Ala Val Val Gly Tyr Val Ser Glu Arg
            180                 185                 190

Phe Pro Gly Arg Arg Ile Ile Leu Val Gly Phe Ser Met Gly Gly Ala
        195                 200                 205

Val Ala Ile Val Glu Gly Ala Gly Asp Pro Arg Val Tyr Ala Val Ala
    210                 215                 220

Ala Asp Ser Pro Tyr Tyr Arg Leu Arg Asp Val Ile Pro Arg Trp Leu
225                 230                 235                 240


Glu Tyr Lys Thr Pro Leu Pro Gly Trp Val Gly Val Leu Ala Gly Phe
                245                 250                 255

Tyr Gly Arg Leu Met Ala Gly Val Asp Leu Gly Phe Gly Pro Ala Gly
            260                 265                 270

Val Glu Arg Val Asp Lys Pro Leu Leu Val Val Tyr Gly Pro Arg Asp
        275                 280                 285

Pro Leu Val Thr Arg Asp Glu Ala Arg Ser Leu Ala Ser Arg Ser Pro
    290                 295                 300

Cys Gly Arg Leu Val Glu Val Pro Gly Ala Gly His Val Glu Ala Val
305                 310                 315                 320

Asp Val Leu Gly Pro Gly Arg Tyr Ala Asp Met Leu Ile Glu Leu Ala
                325                 330                 335

His Glu Glu Cys Pro Pro Gly Ala Gly Gly
            340                 345
```

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 262 AMINO ACIDS
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

```
Met Pro Tyr Val Arg Asn Gly Gly Val Asn Ile Tyr Tyr Glu Leu Val
 1               5                  10                  15

Asp Gly Pro Glu Pro Pro Ile Val Phe Val His Gly Trp Thr Ala Asn
            20                  25                  30

Met Asn Phe Trp Lys Glu Gln Arg Arg Tyr Phe Ala Gly Arg Asn Met
        35                  40                  45

Met Leu Phe Val Asp Asn Arg Gly His Gly Arg Ser Asp Lys Pro Leu
    50                  55                  60

Gly Tyr Asp Phe Tyr Arg Phe Glu Asn Phe Ile Ser Asp Leu Asp Ala
65                  70                  75                  80

Val Val Arg Glu Thr Gly Val Glu Lys Phe Val Leu Val Gly His Ser
                85                  90                  95

Phe Gly Thr Met Ile Ser Met Lys Tyr Cys Ser Glu Tyr Arg Asn Arg
            100                 105                 110

Val Leu Ala Leu Ile Leu Ile Gly Gly Gly Ser Arg Ile Lys Leu Leu
        115                 120                 125

His Arg Ile Gly Tyr Pro Leu Ala Lys Ile Leu Ala Ser Ile Ala Tyr
    130                 135                 140

Lys Lys Ser Ser Arg Leu Val Ala Asp Leu Ser Phe Gly Lys Asn Ala
145                 150                 155                 160

Gly Glu Leu Lys Glu Trp Gly Trp Lys Gln Ala Met Asp Tyr Thr Pro
                165                 170                 175

Ser Tyr Val Ala Met Tyr Thr Tyr Arg Thr Leu Thr Lys Val Asn Leu
            180                 185                 190

Glu Asn Ile Leu Glu Lys Ile Asp Cys Pro Thr Leu Ile Ile Val Gly
        195                 200                 205

Glu Glu Asp Ala Leu Leu Pro Val Ser Lys Ser Val Glu Leu Ser Arg
    210                 215                 220

Arg Ile Glu Asn Ser Lys Leu Val Ile Ile Pro Asn Ser Gly His Cys
225                 230                 235                 240

Val Met Leu Glu Ser Pro Ser Glu Val Asn Arg Ala Met Asp Glu Phe
                245                 250                 255

Ile Ser Ser Ala Gln Phe
            260
```

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 251 AMINO ACIDS
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

```
Leu Arg Leu Arg Lys Phe Glu Glu Ile Asn Leu Val Leu Ser Gly Gly
1               5                   10                  15

Ala Ala Lys Gly Ile Ala His Ile Gly Val Leu Lys Ala Ile Asn Glu
            20                  25                  30

Leu Gly Ile Arg Val Arg Ala Leu Ser Gly Val Ser Ala Gly Ala Ile
        35                  40                  45

Val Ser Val Phe Tyr Ala Ser Gly Tyr Ser Pro Glu Gly Met Phe Ser
    50                  55                  60

Leu Leu Lys Arg Val Asn Trp Leu Lys Leu Phe Lys Phe Lys Pro Pro
65                  70                  75                  80

Leu Lys Gly Leu Ile Gly Trp Glu Lys Ala Ile Arg Phe Leu Glu Glu
                85                  90                  95

Val Leu Pro Tyr Arg Arg Ile Glu Lys Leu Glu Ile Pro Thr Tyr Ile
            100                 105                 110

Cys Ala Thr Asp Leu Tyr Ser Gly Arg Ala Leu Tyr Leu Ser Glu Gly
        115                 120                 125

Ser Leu Ile Pro Ala Leu Leu Gly Ser Cys Ala Ile Pro Gly Ile Phe
    130                 135                 140

Glu Pro Val Glu Tyr Lys Asn Tyr Leu Leu Val Asp Gly Gly Ile Val
145                 150                 155                 160

Asn Asn Leu Pro Val Glu Pro Phe Gln Glu Ser Gly Ile Pro Thr Val
                165                 170                 175

Cys Val Asp Val Leu Pro Ile Glu Pro Glu Lys Asp Ile Lys Asn Ile
            180                 185                 190

Leu His Ile Leu Leu Arg Ser Phe Phe Leu Ala Val Arg Ser Asn Ser
        195                 200                 205

Glu Lys Arg Lys Glu Phe Cys Asp Leu Val Ile Val Pro Glu Leu Glu
    210                 215                 220

Glu Phe Thr Pro Leu Asp Val Arg Lys Ala Asp Gln Ile Met Glu Arg
225                 230                 235                 240

Gly Tyr Ile Lys Ala Leu Glu Val Leu Ser Glu
                245                 250
```

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 297 AMINO ACIDS
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

Met Phe Asn Ile Asn Val Phe Val Asn Ile Ser Trp Leu Tyr Phe Ser
1 5 10 15

Gly Ile Val Met Lys Thr Val Glu Glu Tyr Ala Leu Leu Glu Thr Gly
20 25 30

Val Arg Val Phe Tyr Arg Cys Val Ile Pro Glu Lys Ala Phe Asn Thr
35 40 45

Leu Ile Ile Gly Ser His Gly Leu Gly Ala His Ser Gly Ile Tyr Ile
50 55 60

Ser Val Ala Glu Glu Phe Ala Arg His Gly Phe Gly Phe Cys Met His
65 70 75 80

Asp Gln Arg Gly His Gly Arg Thr Ala Ser Asp Arg Glu Arg Gly Tyr
85 90 95

Val Glu Gly Phe His Asn Phe Ile Glu Asp Met Lys Ala Phe Ser Asp
100 105 110

Tyr Ala Lys Trp Arg Val Gly Gly Asp Glu Ile Ile Leu Leu Gly His
115 120 125

Ser Met Gly Gly Leu Ile Ala Leu Leu Thr Val Ala Thr Tyr Lys Glu
130 135 140

Ile Ala Lys Gly Val Ile Ala Leu Ala Pro Ala Leu Gln Ile Pro Leu
145 150 155 160

Thr Pro Ala Arg Arg Leu Val Leu Ser Leu Ala Ser Arg Leu Ala Pro
165 170 175

His Ser Lys Ile Thr Leu Gln Arg Arg Leu Pro Gln Lys Pro Glu Gly
180 185 190

Phe Gln Arg Ala Lys Asp Ile Glu Tyr Ser Leu Ser Glu Ile Ser Val
195 200 205

Lys Leu Val Asp Glu Met Ile Lys Ala Ser Ser Met Phe Trp Thr Ile
210 215 220

Ala Gly Glu Ile Asn Thr Pro Val Leu Leu Ile His Gly Glu Lys Asp
225 230 235 240

Asn Val Ile Pro Pro Glu Ala Ser Lys Lys Ala Tyr Gln Leu Ile Pro
245 250 255

Ser Phe Pro Lys Glu Leu Lys Ile Tyr Pro Asp Leu Gly His Asn Leu
260 265 270

Phe Phe Glu Pro Gly Ala Val Lys Ile Val Thr Asp Ile Val Glu Trp
275 280 285

Val Lys Asn Leu Pro Arg Glu Asn Pro
290 295

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 262 AMINO ACIDS
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

```
Met Glu Val Tyr Lys Ala Lys Phe Gly Glu Ala Lys Leu Gly Trp Val
 1               5                  10                  15

Val Leu Val His Gly Leu Gly Glu His Ser Gly Arg Tyr Gly Arg Leu
            20                  25                  30

Ile Lys Glu Leu Asn Tyr Ala Gly Phe Gly Val Tyr Thr Phe Asp Trp
        35                  40                  45

Pro Gly His Gly Lys Ser Pro Gly Lys Arg Gly His Thr Ser Val Glu
    50                  55                  60

Glu Ala Met Glu Ile Ile Asp Ser Ile Ile Glu Glu Ile Arg Glu Lys
65                  70                  75                  80

Pro Phe Leu Phe Gly His Ser Leu Gly Gly Leu Thr Val Ile Arg Tyr
                85                  90                  95

Ala Glu Thr Arg Pro Asp Lys Ile Arg Gly Leu Ile Ala Ser Ser Pro
            100                 105                 110

Ala Leu Ala Lys Ser Pro Glu Thr Pro Gly Phe Met Val Ala Leu Ala
        115                 120                 125

Lys Phe Leu Gly Lys Ile Ala Pro Gly Val Val Leu Ser Asn Gly Ile
    130                 135                 140

Lys Pro Glu Leu Leu Ser Arg Asn Arg Asp Ala Val Arg Arg Tyr Val
145                 150                 155                 160

Glu Asp Pro Leu Val His Asp Arg Ile Ser Ala Lys Leu Gly Arg Ser
                165                 170                 175

Ile Phe Val Asn Met Glu Leu Ala His Arg Glu Ala Asp Lys Ile Lys
            180                 185                 190

Val Pro Ile Leu Leu Leu Ile Gly Thr Gly Asp Val Ile Thr Pro Pro
        195                 200                 205

Glu Gly Ser Arg Arg Leu Phe Glu Glu Leu Ala Val Glu Asn Lys Thr
    210                 215                 220

Leu Arg Glu Phe Glu Gly Ala Tyr His Glu Ile Phe Glu Asp Pro Glu
225                 230                 235                 240

Trp Ala Glu Glu Phe His Glu Thr Ile Val Lys Trp Leu Val Glu Lys
                245                 250                 255

Ser Tyr Ser Ser Ala Gln
            260
```

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 249 AMINO ACIDS
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

```
Leu Ile Gly Asn Leu Lys Leu Lys Arg Phe Glu Glu Val Asn Leu Val
1               5                   10                  15

Leu Ser Gly Gly Ala Ala Lys Gly Ile Ala His Ile Gly Val Leu Lys
            20                  25                  30

Ala Leu Glu Glu Leu Gly Ile Lys Val Lys Arg Leu Ser Gly Val Ser
        35                  40                  45

Ala Gly Ala Ile Val Ser Val Phe Tyr Ala Ser Gly Tyr Thr Pro Asp
    50                  55                  60

Glu Met Leu Lys Leu Leu Lys Glu Val Asn Trp Leu Lys Leu Phe Lys
65                  70                  75                  80

Phe Lys Thr Pro Lys Met Gly Leu Met Gly Trp Glu Lys Ala Ala Glu
                85                  90                  95

Phe Leu Glu Lys Glu Leu Gly Val Lys Arg Leu Glu Asp Leu Asn Ile
            100                 105                 110

Pro Thr Tyr Leu Cys Ser Ala Asp Leu Tyr Thr Gly Lys Ala Leu Tyr
        115                 120                 125

Phe Gly Arg Gly Asp Leu Ile Pro Val Leu Leu Gly Ser Lys Ser Ile
    130                 135                 140

Pro Gly Ile Phe Glu Pro Val Glu Tyr Glu Asn Phe Leu Leu Val Asp
145                 150                 155                 160

Gly Gly Ile Val Asn Asn Leu Pro Val Glu Pro Leu Glu Lys Phe Lys
                165                 170                 175

Glu Pro Ile Ile Gly Val Asp Val Leu Pro Ile Thr Gln Glu Arg Lys
            180                 185                 190

Ile Lye Asn Ile Leu His Ile Leu Ile Arg Ser Phe Phe Leu Ala Val
        195                 200                 205

Arg Ser Asn Ser Glu Lys Arg Lys Glu Phe Cys Asn Val Val Ile Glu
    210                 215                 220

Pro Pro Leu Glu Glu Phe Ser Pro Leu Asp Val Asn Lys Ala Asp Glu
225                 230                 235                 240

Ile Phe Cys Gly Asp Met Arg Ala Leu
                245
```

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 339 AMINO ACIDS
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

```
Met Pro Ala Asn Asp Ser Pro Thr Ile Asp Phe Asn Pro Arg Gly Ile
1               5                   10                  15

Leu Arg Asn Ala His Ala Gln Val Ile Leu Ala Thr Ser Gly Leu Arg
            20                  25                  30

Lys Ala Phe Leu Lys Arg Thr His Lys Ser Tyr Leu Ser Thr Ala Gln
        35                  40                  45

Trp Leu Glu Leu Asp Ala Gly Asn Gly Val Thr Leu Ala Gly Glu Leu
    50                  55                  60

Asn Thr Ala Pro Ala Thr Ala Ser Ser Ser His Pro Ala His Lys Asn
65                  70                  75                  80

Thr Leu Val Ile Val Leu His Gly Trp Glu Gly Ser Ser Gln Ser Ala
                85                  90                  95

Tyr Ala Thr Ser Ala Gly Ser Thr Leu Phe Asp Asn Gly Phe Asp Thr
            100                 105                 110

Phe Arg Leu Asn Phe Arg Asp His Gly Asp Thr Tyr His Leu Asn Arg
        115                 120                 125

Gly Ile Phe Asn Ser Ser Leu Ile Asp Glu Val Val Gly Ala Val Lys
    130                 135                 140

Ala Ile Gln Gln Gln Thr Asp Tyr Asp Lys Tyr Cys Leu Met Gly Phe
145                 150                 155                 160

Ser Leu Gly Gly Asn Phe Ala Leu Arg Val Ala Val Arg Glu Gln His
                165                 170                 175

Leu Ala Lys Pro Leu Ala Gly Val Leu Ala Val Cys Pro Val Leu Asp
        180                 185                 190

Pro Ala His Thr Met Met Ala Leu Asn Arg Gly Ala Phe Phe Tyr Gly
    195                 200                 205

Arg Tyr Phe Ala His Lys Trp Lys Arg Ser Leu Thr Ala Lys Leu Ala
210                 215                 220                 225

Ala Phe Pro Asp Tyr Lys Tyr Gly Lys Asp Leu Lys Ser Ile His Thr
                230                 235                 240

Leu Asp Glu Leu Asn Asn Tyr Phe Ile Pro Arg Tyr Thr Gly Phe Asn
            245                 250                 255
```

Ser Val Ser Glu Tyr Phe Lys Ser Tyr Thr Leu Thr Gly Gln Lys Leu
260 265 270

Ala Phe Leu Asn Cys Pro Ser Tyr Ile Leu Ala Ala Gly Asp Asp Pro
275 280 285

Ile Ile Pro Ala Ser Asp Phe Gln Lys Ile Ala Lys Pro Ala Asn Leu
290 295 300 305

His Ile Thr Val Thr Gln Gln Gly Ser His Cys Ala Tyr Leu Glu Asn
310 315 320

Leu His Lys Pro Ser Ala Ala Asp Lys Tyr Ala Val Lys Leu Phe Gly
325 330 335

Ala Cys

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 311 AMINO ACIDS
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

```
Met Leu Asp Met Pro Ile Asp Pro Val Tyr Tyr Gln Leu Ala Glu Tyr
1               5                   10                  15

Phe Asp Ser Leu Pro Lys Phe Asp Gln Phe Ser Ser Ala Arg Glu Tyr
            20                  25                  30

Arg Glu Ala Ile Asn Arg Ile Tyr Glu Glu Arg Asn Arg Gln Leu Ser
        35                  40                  45

Gln His Glu Arg Val Glu Arg Val Glu Asp Arg Thr Ile Lys Gly Arg
    50                  55                  60

Asn Gly Asp Ile Arg Val Arg Val Tyr Gln Gln Lys Pro Asp Ser Pro
65                  70                  75                  80

Val Leu Val Tyr Tyr His Gly Gly Gly Phe Val Ile Cys Ser Ile Glu
                85                  90                  95

Ser His Asp Ala Leu Cys Arg Arg Ile Ala Arg Leu Ser Asn Ser Thr
            100                 105                 110

Val Val Ser Val Asp Tyr Arg Leu Ala Pro Glu His Lys Phe Pro Ala
        115                 120                 125

Ala Val Tyr Asp Cys Tyr Asp Ala Thr Lys Trp Val Ala Glu Asn Ala
    130                 135                 140

Glu Glu Leu Arg Ile Asp Pro Ser Lys Ile Phe Val Gly Gly Asp Ser
145                 150                 155                 160

Ala Gly Gly Asn Leu Ala Ala Ala Val Ser Ile Met Ala Arg Asp Ser
                165                 170                 175

Gly Glu Asp Phe Ile Lys His Gln Ile Leu Ile Tyr Pro Val Val Asn
            180                 185                 190

Phe Val Ala Pro Thr Pro Ser Leu Leu Glu Phe Gly Glu Gly Leu Trp



        195                 200                 205

Ile Leu Asp Gln Lys Ile Met Ser Trp Phe Ser Glu Gln Tyr Phe Ser
    210                 215                 230

Arg Glu Glu Asp Lys Phe Asn Pro Leu Ala Ser Val Ile Phe Ala Asp
235                 240                 245                 250

Leu Glu Asn Leu Pro Pro Ala Leu Ile Ile Thr Ala Glu Tyr Asp Pro
                255                 260                 265

Leu Arg Asp Glu Gly Glu Val Phe Gly Gln Met Leu Arg Arg Ala Gly
            270                 275                 280

Val Glu Ala Ser Ile Val Arg Tyr Arg Gly Val Leu His Gly Phe Ile
        285                 290                 295

Asn Tyr Tyr Pro Val Leu Lys Ala Ala Arg Asp Ala Ile Asn Gln Ile
    300                 305                 310

Ala Ala Leu Leu Val Phe Asp
315                 320
```

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 305 AMINO ACIDS
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(xi) SEQUENCE DESCRIPTION: SEO ID NO:42:

```
Met Pro Leu Asp Pro Arg Ile Lys Lys Leu Leu Glu Ser Ala Leu Thr
                5                   10                  15

Ile Pro Ile Gly Lys Ala Pro Val Glu Glu Val Arg Lys Ile Phe Arg
            20                  25                  30

Gln Leu Ala Ser Ala Ala Pro Lys Val Glu Val Gly Lys Val Glu Asp
        35                  40                  45

Ile Lys Ile Pro Gly Ser Glu Thr Val Ile Asn Ala Arg Val Tyr Phe
    50                  55                  60

Pro Lys Ser Ser Gly Pro Tyr Gly Val Leu Val Tyr Leu His Gly Gly
65                  70                  75                  80

Gly Phe Val Ile Gly Asp Val Glu Ser Tyr Asp Pro Leu Cys Arg Ala
                85                  90                  95

Ile Thr Asn Ala Cys Asn Cys Val Val Val Ser Val Asp Tyr Arg Leu
            100                 105                 110

Ala Pro Glu Tyr Lys Phe Pro Ser Ala Val Ile Asp Ser Phe Asp Ala
        115                 120                 125

Thr Asn Trp Val Tyr Asn Asn Leu Asp Lys Phe Asp Gly Lys Met Gly
    130                 135                 140

Val Ala Ile Ala Gly Asp Ser Ale Gly Gly Asn Leu Ala Ala Val Val
145                 150                 155                 160
```

```
Ala Leu Leu Ser Lys Gly Lys Ile Asn Leu Lys Tyr Gln Ile Leu Val
                165                 170                 175

Tyr Pro Ala Val Ser Leu Asp Asn Val Ser Arg Ser Met Ile Glu Tyr
            180                 185                 190

Ser Asp Gly Phe Phe Leu Thr Arg Glu His Ile Glu Trp Phe Gly Ser
        195                 200                 205

Gln Tyr Leu Arg Ser Pro Ala Asp Leu Leu Asp Phe Arg Phe Ser Pro
    210                 215                 220

Ile Leu Ala Gln Asp Phe Asn Gly Leu Pro Pro Ala Leu Ile Ile Thr
225                 230                 235                 240

Ala Glu Tyr Asp Pro Leu Arg Asp Gln Gly Glu Ala Tyr Ala Asn Lys
                245                 250                 255

Leu Leu Gln Ala Gly Val Ser Val Thr Ser Val Arg Phe Asn Asn Val
            260                 265                 270

Ile His Gly Phe Leu Ser Phe Phe Pro Leu Met Glu Gln Gly Arg Asp
        275                 280                 285

Ala Ile Gly Leu Ile Gly Ser Val Leu Arg Arg Val Phe Tyr Asp Lys
    290                 295                 300

Ile
305
```

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 605 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: GENOMIC DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

```
ATG AAG GTT AAA CAC GTT ATT GTT TTA CAT GGC TTA TAT ATG TCT GGC       48
Met Lys Val Lys His Val Ile Val Leu His Gly Leu Tyr Met Ser Gly
 1               5                  10                  15

TTG GTG ATG CGC CCG TTA TGT TCG CGT CTA GAA GAG TCG GGG GTT AAA       96
Leu Val Met Arg Pro Leu Cys Ser Arg Leu Glu Glu Ser Gly Val Lys
                 20                  25                  30

GTT TTA AAC TTA ACC TAC AAT ACT CGA GAC CCT AAT CGA GAT GCT ATT      144
Val Leu Asn Leu Thr Tyr Asn Thr Arg Asp Pro Asn Arg Asp Ala Ile
    --   35                  40                  45

TTT ACG CAA ATA GAT GAG TTT ATT AGC AAT GAG CCT TCT GCT TTA GTG      192
Phe Thr Gln Ile Asp Glu Phe Ile Ser Asn Glu Pro Ser Ala Leu Val
     50                  55                  60
TGT CAC TCT ATG GGG GGC TTA GTT GCT CGC GCC TAT TTA GAG GCA AAC      240
Cys His Ser Met Gly Gly Leu Val Ala Arg Ala Tyr Leu Glu Ala Asn
 65                  70                  75                  80

TCA GCG CCA AGT CAT CAT GTT GAA AAG GTA ATC ACC TTA GGA ACG CCA      288
Ser Ala Pro Ser His His Val Glu Lys Val Ile Thr Leu Gly Thr Pro
                 85                  90                  95

CAT ACT GGC AGC CAT ATT GCT GAA AAA ATG CAG CAA AAA GGG TTC GAG      336
His Thr Gly Ser His Ile Ala Glu Lys Met Gln Gln Lys Gly Phe Glu
             100                 105                 110
```

```
CTA TTA TTA AAA AAT AGC GTT GAG TTT TTA CTC TCT AAG AAT GGT GAT      384
Leu Leu Leu Lys Asn Ser Val Glu Phe Leu Leu Ser Lys Asn Gly Asp
         115                 120                 125

TGG CCT TTT AAA GCC AAG CTA TAT AGC ATT GCC GGC GAC TTA CCG ATT      432
Trp Pro Phe Lys Ala Lys Leu Tyr Ser Ile Ala Gly Asp Leu Pro Ile
     130                 135                 140

GGC TTA ATG CCA CTC ATT GTA AAA GGC AGC CGC TCT GAT GGC ACT GTA      480
Gly Leu Met Pro Leu Ile Val Lys Gly Ser Arg Ser Asp Gly Thr Val
 145                 150                 155                 160

TTG CTA GAT GAA ACC AAG CTA AAG GGT ATG GCT GAA CAC AAG GTG TTT      528
Leu Leu Asp Glu Thr Lys Leu Lys Gly Met Ala Glu His Lys Val Phe
                 165                 170                 175

CAT TTA AGC CAT ACA AGT ATG ATT TAC TCT CGC CAA GTC GTT AAT TAT      576
His Leu Ser His Thr Ser Met Ile Tyr Ser Arg Gln Val Val Asn Tyr
             180                 185                 190

ATT CTT GAG CGC TTG AAC GAG GAC ATT TA                               605
Ile Leu Glu Arg Leu Asn Glu Asp Ile
         195                 200
```

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 779 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: GENOMIC DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

```
ATG ATA AAA AAC TTC GAC AGA GAA AAT TCT AGC TTA GTA CTG TCC GGT       48
Met Ile Lys Asn Phe Asp Arg Glu Asn Ser Ser Leu Val Leu Ser Gly
 1               5                  10                  15

GGT GGT GCT CTG GGT ATT GCT CAC TTG GGT GTA CTG CAT GAC CTT GAA       96
Gly Gly Ala Leu Gly Ile Ala His Leu Gly Val Leu His Asp Leu Glu
             20                  25                  30

AAA CAA AAT ATT GTA CCA AAT GAA ATT GTT GGT ACA AGT ATG GGT GGT      144
Lys Gln Asn Ile Val Pro Asn Glu Ile Val Gly Thr Ser Met Gly Gly
         35                  40                  45

ATC ATT GGT GCA TCT ATG GCT ATC GGG ATG AAA GAG AAA GAA ATA CTC      192
Ile Ile Gly Ala Ser Met Ala Ile Gly Met Lys Glu Lys Glu Ile Leu
     50                  55                  60

GAA GAA ATC AAA AAC TTT TCC AAT GTC TTC AAC TGG ATA AAA TTC TCT      240
Glu Glu Ile Lys Asn Phe Ser Asn Val Phe Asn Trp Ile Lys Phe Ser
 65                  70                  75                  80

TTT TCC GGT AAT TCT GTT GTC GAT AAC GAG AAG ATC GCT AAG ATA TTT      288
Phe Ser Gly Asn Ser Val Val Asp Asn Glu Lys Ile Ala Lys Ile Phe
                 85                  90                  95

GAT ACT CTT TTT AAA GAC AGA AAG ATG ACA GAT ACG GTG ATC CCT CTT      336
Asp Thr Leu Phe Lys Asp Arg Lys Met Thr Asp Thr Val Ile Pro Leu
            100                 105                 110

AAA CTC ATC GCT ACA AAC TTA CAT AAT GGA CAT AAA AAA GTA TTT ACT      384
Lys Leu Ile Ala Thr Asn Leu His Asn Gly His Lys Lys Val Phe Thr
        115                 120                 125

GCT TCG GAT GAT GTA CTG ATC AAA GAT GCA ATA CTC TCA ACA ATG GCA      432
Ala Ser Asp Asp Val Leu Ile Lys Asp Ala Ile Leu Ser Thr Met Ala
    130                 135                 140

ATA CCC GGT GTA TTT GAA GAA CAT ATT ATT GAT GGT GAA ACC TAT GGC      480
Ile Pro Gly Val Phe Glu Glu His Ile Ile Asp Gly Glu Thr Tyr Gly
145                 150                 155                 160

GAC GGT TTT CTT TGT GAA AAC CTT GGT GTG AAT GAG GCA ACA TTC AAT      528
Asp Gly Phe Leu Cys Glu Asn Leu Gly Val Asn Glu Ala Thr Phe Asn
                165                 170                 175

GAT GTT TTA GCT GTA GAT GTC ATG GGT GAG AAC TCT TTT GAA AAA GCA      576
Asp Val Leu Ala Val Asp Val Met Gly Glu Asn Ser Phe Glu Lys Ala
            180                 185                 190

ATG CCG GAC AAC TTC TTT AAA ACA TCA AAT GTT TTA GAA ATG TTT GAA      624
Met Pro Asp Asn Phe Phe Lys Thr Ser Asn Val Leu Glu Met Phe Glu
        195                 200                 205

AAA TCA ATG CGA CTT TTT ATT TAC AAC CAG ACA CAG ACA CAT ATT AAA      672
Lys Ser Met Arg Leu Phe Ile Tyr Asn Gln Thr Gln Thr His Ile Lys
    210                 215                 220

AAT GCA AAT AAA AAT ATT TAT CTT ATT GAA CCC GTT ACC AAA GAG TAT      720
Asn Ala Asn Lys Asn Ile Tyr Leu Ile Glu Pro Val Thr Lys Glu Tyr
225                 230                 235                 240

AAA ACA TTT CAA TTT CAT AAA CAT AAA GAG ATA CGT GCT TTA GGC TTG      768
Lys Thr Phe Gln Phe His Lys His Lys Glu Ile Arg Ala Leu Gly Leu
                245                 250                 255

GGT TTA CTG TG                                                       779
Gly Leu Leu
```

(2) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 905 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: GENOMIC DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

```
ATG CCC CTA CAT CCA AAG GTA AAG AAA TTA CTT TCC CAG CTA CCT CCC      48
Met Pro Leu His Pro Lys Val Lys Lys Leu Leu Ser Gln Leu Pro Pro
 1               5                   10                  15

CAG GAC TTC TCC AGA AAC GTG CAG GAC CTG AGG AAG GCC TGG GAT TTA      96
Gln Asp Phe Ser Arg Asn Val Gln Asp Leu Arg Lys Ala Trp Asp Leu
             20                  25                  30

CCC TTC TCA GGG AGG AGG GAG ACC CTG AAG AGG GTT GAG GAC CTT GAG     144
Pro Phe Ser Gly Arg Arg Glu Thr Leu Lys Arg Val Glu Asp Leu Glu
         35                  40                  45

ATA CCC ACT AGG GAC GCA CGA ATC AGG GCC AGG GTC TAC ACC CCC TCA     192
Ile Pro Thr Arg Asp Ala Arg Ile Arg Ala Arg Val Tyr Thr Pro Ser
     50                  55                  60

AGT AAG GAA AAC TTA CCC GTC CTT GTT TAC TAT CAC GGC GGT GGC TTC     240
Ser Lys Glu Asn Leu Pro Val Leu Val Tyr Tyr His Gly Gly Gly Phe
 65                  70                  75                  80

GTG TTC GGT AGC GTT GAC AGC TAC GAC GGC CTC GCA TCC CTT ATT GCC     288
Val Phe Gly Ser Val Asp Ser Tyr Asp Gly Leu Ala Ser Leu Ile Ala
                 85                  90                  95

AAG GAA TCT GGG ATT GCG GTT ATC TCC GTG GAG TAT AGG CTC GCC CCT     336
Lys Glu Ser Gly Ile Ala Val Ile Ser Val Glu Tyr Arg Leu Ala Pro
            100                 105                 110

GAG CAC AAG TTC CCC ACC GCA GTC AAC GAC TCG TGG GAT GCG CTT CTC     384
Glu His Lys Phe Pro Thr Ala Val Asn Asp Ser Trp Asp Ala Leu Leu
        115                 120                 125

TGG ATC GCG GAG AAC GGA GGC AAG CTG GGG CTC GAC ACC TCG AGA CTT     432
Trp Ile Ala Glu Asn Gly Gly Lys Leu Gly Leu Asp Thr Ser Arg Leu
    130                 135                 140

GCC GTG GCT GGG GAT AGT GCT GGA GGA AAC CTG TCT GCC GTG GTG TCC     480
Ala Val Ala Gly Asp Ser Ala Gly Gly Asn Leu Ser Ala Val Val Ser
145                 150                 155                 160
```

```
CTC CTG GAC AGG GAC CAG GGT AAG GGA CTG GTT AGT TAT CAG GTC CTA       528
Leu Leu Asp Arg Asp Gln Gly Lys Gly Leu Val Ser Tyr Gln Val Leu
                165                 170                 175

ATC TAC CCA GCA GTG AAC ATG GTC GAT AAC TCC CCA TCC GTC AGG GAG       576
Ile Tyr Pro Ala Val Asn Met Val Asp Asn Ser Pro Ser Val Arg Glu
            180                 185                 190

TAC GGC GAG GGA TAC TTC CTC ACC AGG TCC ATG ATG AAC TGG TTC GGG       624
Tyr Gly Glu Gly Tyr Phe Leu Thr Arg Ser Met Met Asn Trp Phe Gly
        195                 200                 205

ACC ATG TAC TTC TCC TCT GGA AGG GAA GCG GTA TCC CCC TAC GCC TCT       672
Thr Met Tyr Phe Ser Ser Gly Arg Glu Ala Val Ser Pro Tyr Ala Ser
    210                 215                 220

CCA GCC TTG GCT GAC CTA CAT AAC CTC CCA CCC TCA CTG GTG ATC ACT       720
Pro Ala Leu Ala Asp Leu His Asn Leu Pro Pro Ser Leu Val Ile Thr
225                 230                 235                 240

GCA GAG TAT GAT CCC CTA AGG GAT CAG GGA GAG ACC TAC TCT CAC TCC       768
Ala Glu Tyr Asp Pro Leu Arg Asp Gln Gly Glu Thr Tyr Ser His Ser
                245                 250                 255

CTA AAC GAG GCT GGA AAC GTA TCA ACC TTG GTT AGA TAT CAA GGA ATG       816
Leu Asn Glu Ala Gly Asn Val Ser Thr Leu Val Arg Tyr Gln Gly Met
            260                 265                 270

ATT CAC GGC TTC CTG TCC TTC TAC GAG TGG ATA ACT GCC GGT AAA CTA       864
Ile His Gly Phe Leu Ser Phe Tyr Glu Trp Ile Thr Ala Gly Lys Leu
        275                 280                 285

GCC ATT CAC CAC ATT GCT GGG GTT CTG AGA TCT GTC CTT TA                905
Ala Ile His His Ile Ala Gly Val Lue Arg Ser Val Leu Arg Ser Val
    290                 295                 300

Leu                                                                               301
```

(2) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 978 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: GENOMIC DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

```
GTG GCC TTC TTC GAT ATG CCC CTT GAG GAA CTG AAA AAG TAC CGG CCT       48
Val Ala Phe Phe Asp Met Pro Leu Glu Glu Leu Lys Lys Tyr Arg Pro
 1               5                   10                  15

GAA AGG TAC GAG GAG AAA GAT TTC GAT GAG TTC TGG AGG GAA ACA CTT       96
Glu Arg Tyr Glu Glu Lys Asp Phe Asp Glu Phe Trp Arg Glu Thr Leu
             20                  25                  30

AAA GAA AGC GAA GGA TTC CCT CTG GAT CCC GTC TTT GAA AAG GTG GAC      144
Lys Glu Ser Glu Gly Phe Pro Leu Asp Pro Val Phe Glu Lys Val Asp
         35                  40                  45

TTT CAT CTC AAA ACG GTT GAA ACG TAC GAT GTT ACT TTC TCT GGA TAC      192
Phe His Leu Lys Thr Val Glu Thr Tyr Asp Val Thr Phe Ser Gly Tyr
     50                  55                  60

AGG GGG CAG AGA ATA AAG GGC TGG CTT CTT GTT CCG AAG TTG GCG GAA      240
Arg Gly Gln Arg Ile Lys Gly Trp Leu Leu Val Pro Lys Leu Ala Glu
 65                  70                  75                  80

GAA AAG CTT CCA TGC GTC GTG CAG TAC ATA GGT TAC AAT GGT GGA AGG      288
Glu Lys Leu Pro Cys Val Val Gln Tyr Ile Gly Tyr Asn Gly Gly Arg
                 85                  90                  95

GGT TTT CCA CAC GAC TGG CTG TTC TGG CCG TCA ATG GGT TAC ATC TGT      336
Gly Phe Pro His Asp Trp Leu Phe Trp Pro Ser Met Gly Tyr Ile Cys
            100                 105                 110
```

```
TTT GTC ATG GAC ACC AGG GGG CAG GGA AGC GGC TGG ATG AAG GGA GAC      384
Phe Val Met Asp Thr Arg Gly Gln Gly Ser Gly Trp Met Lys Gly Asp
        115                 120                 125

ACA CCG GAT TAC CCT GAG GGT CCA GTC GAT CCA CAG TAC CCC GGA TTC      432
Thr Pro Asp Tyr Pro Glu Gly Pro Val Asp Pro Gln Tyr Pro Gly Phe
    130                 135                 140

ATG ACG AGG GGC ATT CTG GAT CCG GGA ACC TAT TAC TAC AGG CGA GTC      480
Met Thr Arg Gly Ile Leu Asp Pro Gly Thr Tyr Tyr Tyr Arg Arg Val
145                 150                 155                 160

TTC GTG GAT GCG GTC AGG GCG GTG GAA GCA GCC ATT TCC TTC CCG AGA      528
Phe Val Asp Ala Val Arg Ala Val Glu Ala Ala Ile Ser Phe Pro Arg
                165                 170                 175

GTG GAT TCC AGG AAG GTG GTG GTG GCC GGA GGC AGT CAG GGT GGG GGA      576
Val Asp Ser Arg Lys Val Val Val Ala Gly Gly Ser Gln Gly Gly Gly
            180                 185                 190

ATC CCC CTT GCG GTG AGT GCC CTG TCG AAC AGG GTG AAG GCT CTG CTC      624
Ile Pro Leu Ala Val Ser Ala Leu Ser Asn Arg Val Lys Ala Leu Leu
        195                 200                 205

TGC GAT GTG CCG TTT CTG TGC CAC TTC AGA AGG GCC GTG CAA CTT GTC      672
Cys Asp Val Pro Phe Leu Cys His Phe Arg Arg Ala Val Gln Leu Val
    210                 215                 220

GAC ACA CAC CCA TAC GTG GAG ATC ACC AAC TTC CTC AAA ACC CAC AGG      720
Asp Thr His Pro Tyr Val Glu Ile Thr Asn Phe Leu Lys Thr His Arg
225                 230                 235                 240

GAC AAA GAG GAG ATT GTT TTC AGA ACA CTT TCC TAC TTC GAT GGT GTG      768
Asp Lys Glu Glu Ile Val Phe Arg Thr Leu Ser Tyr Phe Asp Gly Val
                245                 250                 255

AAC TTT GCA GCA AGG GCA AAG GTG CCC GCC CTG TTT TCC GTT GGG CTC      816
Asn Phe Ala Ala Arg Ala Lys Val Pro Ala Leu Phe Ser Val Gly Leu
            260                 265                 270

ATG GAC ACC ATC TGT CCT CCC TCG ACG GTC TTC GCC GCT TAC AAC CAC      864
Met Asp Thr Ile Cys Pro Pro Ser Thr Val Phe Ala Ala Tyr Asn His
    275                 280                 285

TAC GCC GGT CCA AAG GAG ATC AGA ATC TAT CCG TAC AAC AAC CAC GAA      912
Tyr Ala Gly Pro Lys Glu Ile Arg Ile Tyr Pro Tyr Asn Asn His Glu
    290                 295                 300

GGT GGA GGT TCT TTC CAG GCA ATT GAG CAG GTG AAA TTC TTG AAG AGA      960
Gly Gly Gly Ser Phe Gln Ala Ile Glu Gln Val Lys Phe Leu Lys Arg
305                 310                 315                 320

CTA TTT GAG GAA GGC TAG                                              978
Leu Phe Glu Glu Gly
                325
```

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 879 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: GENOMIC DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

```
ATG CGC ACC CTC TCC TTC GGT CCG ATG ACC ACA GGG GGA AGC ATT CAC       48
Met Arg Thr Leu Ser Phe Gly Pro Met Thr Thr Gly Gly Ser Ile His
1               5                   10                  15

ATG GCG ACC ATG GAC GTG ATG CGC GGG CCG GGG ATG CAG CGG CTG TCA       96
Met Ala Thr Met Asp Val Met Arg Gly Pro Gly Met Gln Arg Leu Ser
            20                  25                  30

CAG GGC GCC AGG GAG GCC GCG AAC CAC CCC TGG GCG AAG CGA CTG GGC       144
Gln Gly Ala Arg Glu Ala Ala Asn His Pro Trp Ala Lys Arg Leu Gly
        35                  40                  45

CGC ATG GGC TAC GCG GCC AAG GGC GCC GTG TAC GCC ATC ATC GGC GTG       192
Arg Met Gly Tyr Ala Ala Lys Gly Ala Val Tyr Ala Ile Ile Gly Val
    50                  55                  60

CTC GCG CTG AAG CTC GCG GCG GGC GAG GGC GGC CGG ACC ACG GAC AGC       240
Leu Ala Leu Lys Leu Ala Ala Gly Glu Gly Gly Arg Thr Thr Asp Ser
65                  70                  75                  80

CAC GGC GCG GTG AAC ACC GTG GCG CAC GGG CCC TTC GGC GTC GCG CTG       288
His Gly Ala Val Asn Thr Val Ala His Gly Pro Phe Gly Val Ala Leu
                85                  90                  95

CTG GCG GTG CTG GTG GTG GGC CTG CTG GGC TAC GTG GTC TGG AGG TTC       336
Leu Ala Val Leu Val Val Gly Leu Leu Gly Tyr Val Val Trp Arg Phe
            100                 105                 110

GCC CAG GCC TTC GTG GAC ACG GAG GAC AAG GGC TCC GAC GCG AAG GGA       384
Ala Gln Ala Phe Val Asp Thr Glu Asp Lys Gly Ser Asp Ala Lys Gly
        115                 120                 125

ATC GCC ACG CGC GCC ATG TAC TTC CTC AGC GGC TGC ATC TAC GCG TCG       432
Ile Ala Thr Arg Ala Met Tyr Phe Leu Ser Gly Cys Ile Tyr Ala Ser
    130                 135                 140

CTG GCC TTC TTC GCC GCG CAG TCC CTG GTG GGC GCC GCG CAC GGC CGG       480
Leu Ala Phe Phe Ala Ala Gln Ser Leu Val Gly Ala Ala His Gly Arg
145                 150                 155                 160

AGC AAG GGG ACG CAG GGC TGG ACG GCC ACG CTG ATG GAG CAG CCC TTT       528
Ser Lys Gly Thr Gln Gly Trp Thr Ala Thr Leu Met Glu Gln Pro Phe
                165                 170                 175

GGC CGC GTG CTG GTG GCG CTG GTG GGG CTG GGC ATC GTG GGC TTC GCG       576
Gly Arg Val Leu Val Ala Leu Val Gly Leu Gly Ile Val Gly Phe Ala
            180                 185                 190

CTG AAG CAG TTC CAC ACC GCG TGG AAG GCG AAG TTC CGG GAG AAG CTC       624
Leu Lys Gln Phe His Thr Ala Trp Lys Ala Lys Phe Arg Glu Lys Leu
        195                 200                 205

ACC CTC ACC GGA CTG GCT GCC CGG AAG CAG CAC CAC ATC GAG CGC ATG       672
Thr Leu Thr Gly Leu Ala Ala Arg Lys Gln His His Ile Glu Arg Met
    210                 215                 220

TGC CAG TTC GGC ATC GCC GCG CGC GGC GTG GTG TTC GCC GTC ATC GGC       720
Cys Gln Phe Gly Ile Ala Ala Arg Gly Val Val Phe Ala Val Ile Gly
225                 230                 235                 240

GGC TTC CTC GTC CGC TCC GCC GTG GAC GCG AAC CCC GGC GAG GCC AAG       768
Gly Phe Leu Val Arg Ser Ala Val Asp Ala Asn Pro Gly Glu Ala Lys
                245                 250                 255

GGC CTG GGA GAG GCC CTG GCC GTC GTC GCG AGG CAG CCG TCC GGC GAC       816
Gly Leu Gly Glu Ala Leu Ala Val Val Ala Arg Gln Pro Ser Gly Asp
            260                 265                 270

GTG CTC CTG GGG GTG GTG GCG GCG GGC CTG GTG GCC TAC GCC GCC TAC       864
Val Leu Leu Gly Val Val Ala Ala Gly Leu Val Ala Tyr Ala Ala Tyr
        275                 280                 285

CTG TTC CTC CAG GCG CGC TAC CGC GAA CTC TAG                           897
Leu Phe Leu Gln Ala Arg Tyr Arg Glu Leu
    290                 295
```

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 914 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: GENOMIC DNA

(xi) SEQUENCE DESCRIPTION: SEQ-ID NO:48:

```
ATG AGC AAA TTC GCA ATA CTC TGG GCG TTG ATA ACG GCA TAC CTG CCG       48
Met Ser Lys Phe Ala Ile Leu Trp Ala Leu Ile Thr Ala Tyr Leu Pro
1               5                   10                  15

GAA CCT GTG ATG AAA CTG GTA TAT TTA GGG CGG CGC GAA ACG CTT GGG       96
Glu Pro Val Met Lys Leu Val Tyr Leu Gly Arg Arg Glu Thr Leu Gly
            20                  25                  30

GCA CGG ACG CTT GAC GTT AAA GCC CAA GCT GTC GGG CGG CTG GCC AAT      144
Ala Arg Thr Leu Asp Val Lys Ala Gln Ala Val Gly Arg Leu Ala Asn
        35                  40                  45

GCA ACA AGA CCT GTC GGG GTG ATT CCG ACG GTC GAG GAA AGC CGG AAG      192
Ala Thr Arg Pro Val Gly Val Ile Pro Thr Val Glu Glu Ser Arg Lys
    50                  55                  60

ATG ACG GAT AAA GCC GTT AGC CTT TTT GAT CAG CCC GCC CCC GAA TTA      240
Met Thr Asp Lys Ala Val Ser Leu Phe Asp Gln Pro Ala Pro Glu Leu
65                  70                  75                  80

TTC CGT AAA AAA GAC ATT CAG ATT GAC GGG GCT GAA GGG CCT ATT GAT      288
Phe Arg Lys Lys Asp Ile Gln Ile Asp Gly Ala Glu Gly Pro Ile Asp
                85                  90                  95

GCC CGT ATT TAC AGC GGC CCT GCA AAA CAT CGC CCG CGG CCA ATG CTA      336
Ala Arg Ile Tyr Ser Gly Pro Ala Lys His Arg Pro Arg Pro Ile Leu
            100                 105                 110

GTG TAT TTT CAC GGC GGT GGC TGG GTT CAG GGC AAT CTG GAC AGC CAT      384
Val Tyr Phe His Gly Gly Gly Trp Val Gln Gly Asn Leu Asp Ser His
        115                 120                 125

GAC GGG GTT TGC GGC AAG CTG GCA AAA TGG GCG AAC TGC ATT GTT ATC      432
Asp Gly Val Cys Gly Lys Leu Ala Lys Trp Ala Asn Cys Ile Val Ile
    130                 135                 140

TCG GTC GAT TAT CGT CTA GCG CCC GAA CAC AAA TTT CCT TGT GCG CCG      480
Ser Val Asp Tyr Arg Leu Ala Pro Glu His Lys Phe Pro Cys Ala Pro
145                 150                 155                 160

CTT GAT GCG ATT GCG GCC TAT AAA TGG GTG CGC GCC AAC GCA ACA AAC      528
Leu Asp Ala Ile Ala Ala Tyr Lys Trp Val Arg Ala Asn Ala Thr Asn
                165                 170                 175

CTT GGC GGC GAT CCT GAA CGT ATC GGC GTT GGC GGC GAT AGC GCA GGG      576
Leu Gly Gly Asp Pro Glu Arg Ile Gly Val Gly Gly Asp Ser Ala Gly
            180                 185                 190

GGC AAT CTT GCC GCC GTT GTC TGC CAA CAA ACC GCC ATG AAC GGC GAG      624
Gly Asn Leu Ala Ala Val Val Cys Gln Gln Thr Ala Met Asn Gly Glu
        195                 200                 205

CGC ACA CCA GAT CTG CAA GTC CTG ATC TAT CCG GCG CTG GAT GCA CGC      672
Arg Thr Pro Asp Leu Gln Val Leu Ile Tyr Pro Ala Leu Asp Ala Arg
    210                 215                 220

ATG ATC TCG ACC TCG ATG GAG GAA TTG CGT GAT GCC TAC ATC TTG CCG      720
Met Ile Ser Thr Ser Met Glu Glu Leu Arg Asp Ala Tyr Ile Leu Pro
225                 230                 235                 240

AAA TCC AGA ATG GAG TAT TTC CTC GGC CTA TAT ACG CGT GGC CCT GAC      768
Lys Ser-Arg Met Glu Tyr Phe Leu Gly Leu Tyr Thr Arg Gly Pro Asp
                245                 250                 255

GAT ATC GAG GAC CTT AGG ATG TCG CCA ATT CTC AGG GAT ACC GTC GCG      816
Asp Ile Glu Asp Leu Arg Met Ser Pro Ile Leu Arg Asp Thr Val Ala
            260                 265                 270

GAT CAA CCC CAA GCC TGC ATT GTC ACC TGT GGG TTT GAC CCT GCG CGA      864
Asp Gln Pro Gln Ala Cys Ile Val Thr Cys Gly Phe Asp Pro Ala Arg
        275                 280                 285

CGA CGG GAA CAC CTA CGC CGA ACG CTT AAT TGC CGA GGG GAT AGA CGT      912
Arg Arg Glu His Leu Arg Arg Thr Leu Asn Cys Arg Gly Asp Arg Arg
    290                 295                 300

TA                                                                   914
```

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 926 NDCLEOTIDES
(B) TYPE: NUCLSIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: GENOMIC DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

```
GTG AGC ATT CGT CTG CGA CTG TTA AAC TGG TTT TTG AAT ACC TTT GAA         48
Val Ser Ile Arg Leu Arg Leu Leu Asn Trp Phe Leu Asn Thr Phe Glu
 1               5                   10                  15

AAA CCA AAA CTG GCC GCG GCC AAA ACG CCG GAT GAT TTG CGA AAA TCG         96
Lys Pro Lys Leu Ala Ala Ala Lys Thr Pro Asp Asp Leu Arg Lys Ser
             20                  25                  30

TTT GAA TTA AAG GCG AGG TTT TTG TTT CCG GCG CCA CGT AAA ACA AGG        144
Phe Glu Leu Lys Ala Arg Phe Leu Phe Pro Ala Pro Arg Lys Thr Arg
         35                  40                  45

TTT AGT CAT GAT GTA TTG CAG TCA GGC ATC GGG TCG GTA AAT GCC CAG        192
Phe Ser His Asp Val Leu Gln Ser Gly Ile Gly Ser Val Asn Ala Gln
     50                  55                  60

TGG GCG AAA TCC AAA TCT GCA TCT GAT GAC AGG GTA ATC CTG TAT TTT        240
Trp Ala Lys Ser Lys Ser Ala Ser Asp Asp Arg Val Ile Leu Tyr Phe
 65                  70                  75                  80

CAT GGG GGA GGG TAT GTT TTT GGG TCA CCA AAA ACG CAC CGT GCA ATG        288
His Gly Gly Gly Tyr Val Phe Gly Ser Pro Lys Thr His Arg Ala Met
                 85                  90                  95

TTG GCG CGC TTG TCG GCA ATG ACA GGT CTT TCT GCG TGC CTT CCA GAT        336
Leu Ala Arg Leu Ser Ala Met Thr Gly Leu Ser Ala Cys Leu Pro Asp
            100                 105                 110

TAT AGG TTG GCA CCA GAG CAC CCA TTT CCA GCC GCG ATC GAA GAT GCA        384
Tyr Arg Leu Ala Pro Glu His Pro Phe Pro Ala Ala Ile Glu Asp Ala
        115                 120                 125

GTT TTA TCG TAT AAA TGT TTA CTA GAG CGA GCA ATC GAG CCC CAA AAT        432
Val Leu Ser Tyr Lys Cys Leu Leu Glu Arg Ala Ile Glu Pro Gln Asn
    130                 135                 140

ATT ATA CTG GGG GGG GAC AGT GCT GGT GGC GGT TTG GTT CTT GCT TTG        480
Ile Ile Leu Gly Gly Asp Ser Ala Gly Gly Gly Leu Val Leu Ala Leu
145                 150                 155                 160

CTT GCA GAA ATC AAG GCC CAA TCC TTG CCC AAA CCT GCT GGC GTT TTT        528
Leu Ala Glu Ile Lys Ala Gln Ser Leu Pro Lys Pro Ala Gly Val Phe
                165                 170                 175

GCC TTG TCG CCT TTG GTT GAT TTA TCA TTT TCG GGC CTT TCG TTT TCT        576
Ala Leu Ser Pro Leu Val Asp Leu Ser Phe Ser Gly Leu Ser Phe Ser
            180                 185                 190

AAA AAT GCC CAA ACC GAT GTG ATG TTG CCC GCA TCA CGG GCT GCG GAT        624
Lys Asn Ala Gln Thr Asp Val Met Leu Pro Ala Ser Arg Ala Ala Asp
        195                 200                 205

ATG GCG ACC TTG TAT TTG GAT GGG GCC GAT GCA GAT GAT CCA CGT GCA        672
Met Ala Thr Leu Tyr Leu Asp Gly Ala Asp Ala Asp Asp Pro Arg Ala
    210                 215                 220

TCG CCG CTG CAG GCG GAT TTT TCT GGC ATG CCG CCT GTA TTT CTG ACA        720
Ser Pro Leu Gln Ala Asp Phe Ser Gly Met Pro Pro Val Phe Leu Thr
225                 230                 235                 240

GCA AGT GAC AGT GAA ATC CTG TTG GAT GAT TGC CTG CGG ATG GCG GAT        768
Ala Ser Asp Ser Glu Ile Leu Leu Asp Asp Cys Leu Arg Met Ala Asp
                245                 250                 255

CAC TTG CGT GCG CAA GGT GTC GTT GTG ACA GAC CGG ATT GTT GAA AAC        816
His Leu Arg Ala Gln Gly Val Val Val Thr Asp Arg Ile Val Glu Asn
            260                 265                 270
```

```
CAT CCA CAT GTT TGG CAT ATT TTT CAA CGC CTT CTA CCC GAA GCA GAT     864
His Pro His Val Trp His Ile Phe Gln Arg Leu Leu Pro Glu Ala Asp
        275                 280                 285

CAG GGG CTG CGG GCG ATT GCC GCG TGG ATT AAA CCT CTT TTA TCA GGT     912
Gln Gly Leu Arg Ala Ile Ala Ala Trp Ile Lys Pro Leu Leu Ser Gly
    290                 295                 300

TCA AAC GAA AGC TA                                                  926
Ser Asn Glu Ser
305
```

(2) INFORMATION FOR SEQ ID NO:50:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 713 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: GENOMIC DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

```
ATG CTT ACA TTT AAT GTT TTA TAT GGT ATG ATG AAA CAA AAA CTA GCA       48
Met Leu Thr Phe Asn Val Leu Tyr Gly Met Met Lys Gln Lys Leu Ala
1               5                   10                  15

GCA ATT CTC ATG TTT TTA GGG CTA TCA GCA GCA GAG GCT CAA GAC TGG       96
Ala Ile Leu Met Phe Leu Gly Leu Ser Ala Ala Glu Ala Gln Asp Trp
            20                  25                  30

CCT GAC CTA CAG AAA TAT CGT AGT GCT AAT AAA GAA GCC AAA TTA CTT      144
Pro Asp Leu Gln Lys Tyr Arg Ser Ala Asn Lys Glu Ala Lys Leu Leu
        35                  40                  45

CCA AAG GAA AAC CGG AAG GTG GTT TTT ATG GGC AAC TCC ATT ACA GAA      192
Pro Lys Glu Asn Arg Lys Val Val Phe Met Gly Asn Ser Ile Thr Glu
    50                  55                  60

GCC TGG ATT AGT CAG CGA CCT GAG TTT TTT AGT GAA AAT GGG TTT ATC      240
Ala Trp Ile Ser Gln Arg Pro Glu Phe Phe Ser Glu Asn Gly Phe Ile
65                  70                  75                  80

GGT CGA GGC ATC AGT GGC CAG ACA ACC CCT CAG ATG TTG TTG AGA TTC      288
Gly Arg Gly Ile Ser Gly Gln Thr Thr Pro Gln Met Leu Leu Arg Phe
                85                  90                  95

CGA CAG GAT GTG ATA GAC CTG CAG CCA AAG GCT GTA GTG ATA CTA GCT      336
Arg Gln Asp Val Ile Asp Leu Gln Pro Lys Ala Val Val Ile Leu Ala
            100                 105                 110

GGT ACC AAT GAC GTA GCT CAA AAT ACC GGG CCG ATG ACC ATT GAG GAA      384
Gly Thr Asn Asp Val Ala Gln Asn Thr Gly Pro Met Thr Ile Glu Glu
        115                 120                 125

TCG CTT GCT AAC ATT AAG TCT ATG GTG GAG CTG GCG CAA GCC AAT GGG      432
Ser Leu Ala Asn Ile Lys Ser Met Val Glu Leu Ala Gln Ala Asn Gly
    130                 135                 140

ATC ACG CCT GTT TTG TGT ACC GTG CTG CCT GCA GAT CGT TTC AGC TGG      480
Ile Thr Pro Val Leu Cys Thr Val Leu Pro Ala Asp Arg Phe Ser Trp
145                 150                 155                 160

CGA CCT GAG CTT ACA CCC GCA GAA ACT ATC ATT GCC CTC AAT CAG CTC      528
Arg Pro Glu Leu Thr Pro Ala Glu Thr Ile Ile Ala Leu Asn Gln Leu
                165                 170                 175

ATT AAG CAA TAT GCC GAG GCA CAG GGC CTG GCC CTG GTG GAT TAT CAT      576
Ile Lys Gln Tyr Ala Glu Ala Gln Gly Leu Ala Leu Val Asp Tyr His
            180                 185                 190

GCT GCA CTC ACC AAT AAA GGT GGA GGA CTT CCG GTG AAA TAC GGA GAA      624
Ala Ala Leu Thr Asn Lys Gly Gly Gly Leu Pro Val Lys Tyr Gly Glu
        195                 200                 205

GAT GGT GTG CAT CCA AAT GTA GCA GGC TAT CAG GTG ATG GAA AAC ATT      672
Asp Gly Val His Pro Asn Val Ala Gly Tyr Gln Val Met Glu Asn Ile
    210                 215                 220

GTT TTA CCG GTC ATT TCC AGC GAG TTG GCA AAG CTG AAG TA               713
Val Leu Pro Val Ile Ser Ser Glu Leu Ala Lys Leu Lys
225                 230                 235
```

(2) INFORMATION POR SEQ ID NO:51:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 978 NOCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: GENOMIC DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

```
ATG GCC TTC TTC GAT TTA CCA CTC GAA GAA CTG AAG AAA TAT CGT CCA      48
Met Ala Phe Phe Asp Leu Pro Leu Glu Glu Leu Lys Lys Tyr Arg Pro
 1               5                  10                  15

GAG CGG TAC GAA GAG AAA GAC TTC GAT GAG TTC TGG GAA GAG ACA CTC      96
Glu Arg Tyr Glu Glu Lys Asp Phe Asp Glu Phe Trp Glu Glu Thr Leu
            20                  25                  30

GCA GAG AGC GAA AAG TTC CCC TTA GAC CCC GTC TTC GAG AGG ATG GAG     144
Ala Glu Ser Glu Lys Phe Pro Leu Asp Pro Val Phe Glu Arg Met Glu
        35                  40                  45

TCT CAC CTC AAA ACA GTC GAA GCG TAC GAT GTC ACC TTC TCC GGA TAC     192
Ser His Leu Lys Thr Val Glu Ala Tyr Asp Val Thr Phe Ser Gly Tyr
    50                  55                  60

AGG GGA CAG AGG ATC AAA GGG TGG CTC CTT GTT CCA AAA CTG GAA GAA     240
Arg Gly Gln Arg Ile Lys Gly Trp Leu Leu Val Pro Lys Leu Glu Glu
65                  70                  75                  80

GAA AAA CTT CCC TGC GTT GTG CAG TAC ATA GGA TAC AAC GGT GGA AGA     288
Glu Lys Leu Pro Cys Val Val Gln Tyr Ile Gly Tyr Asn Gly Gly Arg
                85                  90                  95

GGA TTC CCT CAC GAC TGG CTG TTC TGG CCT TCT ATG GGT TAC ATA TGT     336
Gly Phe Pro His Asp Trp Leu Phe Trp Pro Ser Met Gly Tyr Ile Cys
           100                 105                 110

TTC GTC ATG GAT ACT CGA GGT CAG GGA AGC GGC TGG CTG AAA GGA GAC     384
Phe Val Met Asp Thr Arg Gly Gln Gly Ser Gly Trp Leu Lys Gly Asp
       115                 120                 125

ACA CCG GAT TAC CCT GAG GGT CCC GTT GAC CCT CAG TAT CCA GGA TTC     432
Thr Pro Asp Tyr Pro Glu Gly Pro Val Asp Pro Gln Tyr Pro Gly Phe
    130                 135                 140

ATG ACA AGA GGA ATA CTG GAT CCC AGA ACT TAC TAC TAC AGA CGA GTC     480
Met Thr Arg Gly Ile Leu Asp Pro Arg Thr Tyr Tyr Tyr Arg Arg Val
145                 150                 155                 160

TTC ACG GAC GCT GTC AGA GCC GTT GAA GCT GCT GCT TCT TTT CCT CAG     528
Phe Thr Asp Ala Val Arg Ala Val Glu Ala Ala Ala Ser Phe Pro Gln
                165                 170                 175

GTA GAT CAA GAA AGA ATC GTG ATA GCT GGA GGC AGT CAG GGT GGC GGA     576
Val Asp Gln Glu Arg Ile Val Ile Ala Gly Gly Ser Gln Gly Gly Gly
            180                 185                 190

ATA GCC CTT GCG GTG AGC GCT CTC TCA AAG AAA GCA AAG GCT CTT CTG     624
Ile Ala Leu Ala Val Ser Ala Leu Ser Lys Lys Ala Lys Ala Leu Leu
        195                 200                 205

TGC GAT GTG CCG TTT CTG TGT CAC TTC AGA AGA GCA GTA CAG CTT GTG     672
Cys Asp Val Pro Phe Leu Cys His Phe Arg Arg Ala Val Gln Leu Val
    210                 215                 220
```

```
GAT ACG CAT CCA TAC GCG GAG ATC ACG AAC TTT CTA AAG ACC CAC AGA       720
Asp Thr His Pro Tyr Ala Glu Ile Thr Asn Phe Leu Lys Thr His Arg
225                 230                 235                 240

GAC AAG GAA GAA ATC GTG TTC AGG ACT CTT TCC TAT TTC GAT GGA GTG       768
Asp Lys Glu Glu Ile Val Phe Arg Thr Leu Ser Tyr Phe Asp Gly Val
                245                 250                 255

AAC TTC GCA GCC AGA GCG AAG ATC CCT GCG CTG TTT TCT GTG GGT CTC       816
Asn Phe Ala Ala Arg Ala Lys Ile Pro Ala Leu Phe Ser Val Gly Leu
            260                 265                 270

ATG GAC AAC ATT TGT CCT CCT TCA ACG GTT TTC GCT GCC TAC AAT TAC       864
Met Asp Asn Ile Cys Pro Pro Ser Thr Val Phe Ala Ala Tyr Asn Tyr
        275                 280                 285

TAC GCT GGA CCG AAG GAA ATC AGA ATC TAT CCG TAC AAC AAC CAC GAG       864
Tyr Ala Gly Pro Lys Glu Ile Arg Ile Tyr Pro Tyr Asn Asn His Glu
    290                 295                 300

GGA GGA GGC TCT TTC CAA GCG GTT GAA CAG GTG AAA TTC TTG AAA AAA       912
Gly Gly Gly Ser Phe Gln Ala Val Glu Gln Val Lys Phe Leu Lys Lys
305                 310                 315                 320

CTA TTT GAG AAA GGC TAA                                               930
Leu Phe Glu Lys Gly
                325
```

(2) INFORMATION FOR SEQ ID NO:52:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 660 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: GENOMIC DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

```
TTG AAG TAC TTC AAA GCC CGG CTT GCC GGC ATC ACC TTG CTC GGC CTG         48
Leu Lys Tyr Phe Lys Ala Arg Leu Ala Gly Ile Thr Leu Leu Gly Leu
 1              5                  10                  15

CTG GCC TGC ACC TCG GCC TCG GCG CAG ACC GAG CCC ATC GTG TTC GTG         96
Leu Ala Cys Thr Ser Ala Ser Ala Gln Thr Glu Pro Ile Val Phe Val
            20                  25                  30

CAC GGC TAT TCC GGC AGC GCA TCC AAC TGG GAC ACC ATG CTG GGC CGC        144
His Gly Tyr Ser Gly Ser Ala Ser Asn Trp Asp Thr Met Leu Gly Arg
        35                  40                  45

TTC CGG TCG AAC GGT TAT GCG TCC GGC TCG CTC TAC ACC TTC AAC TAC        192
Phe Arg Ser Asn Gly Tyr Ala Ser Gly Ser Leu Tyr Thr Phe Asn Tyr
    50                  55                  60

AAC TCG TTG GTC AGC AGC AAC CGC ACC AGC GCC AGC GAG CTG CGC AGC        240
Asn Ser Leu Val Ser Ser Asn Arg Thr Ser Ala Ser Glu Leu Arg Ser
65                  70                  75                  80

TTC GTC AAC ACC GTG CGT TCG CGC CAC GGC AAC GCC CGC ATC GCG CTG        288
Phe Val Asn Thr Val Arg Ser Arg His Gly Asn Ala Arg Ile Ala Leu
                85                  90                  95

GTC GCC CAC TCC AAC GGC GGG CTG GTG TCG CGC TGG TAT CGC GCG GAG        336
Val Ala His Ser Asn Gly Gly Leu Val Ser Arg Trp Tyr Arg Ala Glu
            100                 105                 110

CTG GGC GGC GAA ACG GCC ACC CGC CGC TTC GTG ACG CTG GGC ACG CCG        384
Leu Gly Gly Glu Thr Ala Thr Arg Arg Phe Val Thr Leu Gly Thr Pro
        115                 120                 125

CAC CGG GGC ACC ACC TGG GCC TAT GCG TGC TAC AGC CCC GCA TGT TTC        432
His Arg Gly Thr Thr Trp Ala Tyr Ala Cys Tyr Ser Pro Ala Cys Phe
    130                 135                 140
```

```
GAG ATG CGC CCC GGC TCC AGC TTG CTG ACC ACG CTG GGC TCG CGT GCC        480
Glu Met Arg Pro Gly Ser Ser Leu Leu Thr Thr Leu Gly Ser Arg Ala
145                 150                 155                 160

TGC GAC CGC TCG CTG TGG TCG AAC ACC GAC GGC ATC ATC CTG CCG GCG        528
Cys Asp Arg Ser Leu Trp Ser Asn Thr Asp Gly Ile Ile Leu Pro Ala
                165                 170                 175

TCC AGC GCG CAG TGT GGT GTC AGC ACG CGC ACT GCC GAC GTC AGC CAT        576
Ser Ser Ala Gln Cys Gly Val Ser Thr Arg Thr Ala Asp Val Ser His
            180                 185                 190

CTC GAC CTG CTG ACC GAC TCT CGC GTG TAC ACG CAG TTG CGC ACG CAG        624
Leu Asp Leu Leu Thr Asp Ser Arg Val Tyr Thr Gln Leu Arg Thr Gln
        195                 200                 205

TTG CAA TGA GGG TGA CGG TGC ACC GAA CGT GCA CCT G                      661
Leu Gln End Gly End Arg Cys Thr Glu Arg Ala Pro
    210                 215                 220
```

(2) INFORMATION FOR SEQ ID NO:53:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 201 AMINO ACIDS
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

```
Met Lys Val Lys His Val Ile Val Leu His Gly Leu Tyr Met Ser Gly
1               5                   10                  15

Leu Val Met Arg Pro Leu Cys Ser Arg Leu Glu Glu Ser Gly Val Lys
            20                  25                  30

Val Leu Asn Leu Thr Tyr Asn Thr Arg Asp Pro Asn Arg Asp Ala Ile
        35                  40                  45

Phe Thr Gln Ile Asp Glu Phe Ile Ser Asn Glu Pro Ser Ala Leu Val
    50                  55                  60

Cys His Ser Met Gly Gly Leu Val Ala Arg Ala Tyr Leu Glu Ala Asn
65                  70                  75                      80

Ser Ala Pro Ser His His Val Glu Lys Val Ile Thr Leu Gly Thr Pro
                85                  90                  95

His Thr Gly Ser His Ile Ala Glu Lys Met Gln Gln Lys Gly Phe Glu
            100                 105                 110

Leu Leu Leu Lys Asn Ser Val Glu Phe Leu Leu Ser Lys Asn Gly Asp
        115                 120                 125

Trp Pro Phe Lys Ala Lys Leu Tyr Ser Ile Ala Gly Asp Leu Pro Ile
    130                 135                 140

Gly Leu Met Pro Leu Ile Val Lys Gly Ser Arg Ser Asp Gly Thr Val
145                 150                 155                     160

Leu Leu Asp Glu Thr Lys Leu Lys Gly Met Ala Glu His Lys Val Phe
                165                 170                 175

His Leu Ser His Thr Ser Met Ile Tyr Ser Arg Gln Val Val Asn Tyr
            180                 185                 190

Ile Leu Glu Arg Leu Asn Glu Asp Ile
        195                 200
```

(2) INFORMATION FOR SEQ ID NO:54 :

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 259 AMINO ACIDS
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

```
Met Ile Lys Asn Phe Asp Arg Glu Asn Ser Ser Leu Val Leu Ser Gly
1               5                   10                  15

Gly Gly Ala Leu Gly Ile Ala His Leu Gly Val Leu His Asp Leu Glu
            20                  25                  30

Lys Gln Asn Ile Val Pro Asn Glu Ile Val Gly Thr Ser Met Gly Gly
        35                  40                  45

Ile Ile Gly Ala Ser Met Ala Ile Gly Met Lys Glu Lys Glu Ile Leu
    50                  55                  60

Glu Glu Ile Lys Asn Phe Ser Asn Val Phe Asn Trp Ile Lys Phe Ser
65                  70                  75                  80

Phe Ser Gly Asn Ser Val Val Asp Asn Glu Lys Ile Ala Lys Ile Phe
                85                  90                  95

Asp Thr Leu Phe Lys Asp Arg Lys Met Thr Asp Thr Val Ile Pro Leu
            100                 105                 110

Lys Leu Ile Ala Thr Asn Leu His Asn Gly His Lys Lys Val Phe Thr
        115                 120                 125

Ala Ser Asp Asp Val Leu Ile Lys Asp Ala Ile Leu Ser Thr Met Ala
    130                 135                 140

Ile Pro Gly Val Phe Glu Glu His Ile Ile Asp Gly Glu Thr Tyr Gly
145                 150                 155                 160

Asp Gly Phe Leu Cys Glu Asn Leu Gly Val Asn Glu Ala Thr Phe Asn
                165                 170                 175

Asp Val Leu Ala Val Asp Val Met Gly Glu Asn Ser Phe Glu Lys Ala
            180                 185                 190

Met Pro Asp Asn Phe Phe Lys Thr Ser Asn Val Leu Glu Met Phe Glu
        195                 200                 205

Lys Ser Met Arg Leu Phe Ile Tyr Asn Gln Thr Gln Thr His Ile Lys
    210                 215                 220

Asn Ala Asn Lys Asn Ile Tyr Leu Ile Glu Pro Val Thr Lys Glu Tyr
225                 230                 235                 240

Lys Thr Phe Gln Phe His Lys His Lys Glu Ile Arg Ala Leu Gly Leu
                245                 250                 255

Gly Leu Leu
```

(2) INFORMATION FOR SEQ ID NO:55:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 301 AMINO ACIDS
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

```
Met Pro Leu His Pro Lys Val Lys Lys Leu Leu Ser Gln Leu Pro Pro
1               5                   10                  15

Gln Asp Phe Ser Arg Asn Val Gln Asp Leu Arg Lys Ala Trp Asp Leu
            20                  25                  30

Pro Phe Ser Gly Arg Arg Glu Thr Leu Lys Arg Val Glu Asp Leu Glu
        35                  40                  45

Ile Pro Thr Arg Asp Ala Arg Ile Arg Ala Arg Val Tyr Thr Pro Ser
    50                  55                  60
```

```
Ser Lys Glu Asn Leu Pro Val Leu Val Tyr Tyr His Gly Gly Gly Phe
65                  70                  75                  80

Val Phe Gly Ser Val Asp Ser Tyr Asp Gly Leu Ala Ser Leu Ile Ala
                85                  90                  95

Lys Glu Ser Gly Ile Ala Val Ile Ser Val Glu Tyr Arg Leu Ala Pro
            100                 105                 110

Glu His Lys Phe Pro Thr Ala Val Asn Asp Ser Trp Asp Ala Leu Leu
        115                 120                 125

Trp Ile Ala Glu Asn Gly Gly Lys Leu Gly Leu Asp Thr Ser Arg Leu
    130                 135                 140

Ala Val Ala Gly Asp Ser Ala Gly Gly Asn Leu Ser Ala Val Val Ser
145                 150                 155                 160

Leu Leu Asp Arg Asp Gln Gly Lys Gly Leu Val Ser Tyr Gln Val Leu
                165                 170                 175

Ile Tyr Pro Ala Val Asn Met Val Asp Asn Ser Pro Ser Val Arg Glu
            180                 185                 190

Tyr Gly Glu Gly Tyr Phe Leu Thr Arg Ser Met Met Asn Trp Phe Gly
        195                 200                 205

Thr Met Tyr Phe Ser Ser Gly Arg Glu Ala Val Ser Pro Tyr Ala Ser
    210                 215                 220

Pro Ala Leu Ala Asp Leu His Asn Leu Pro Pro Ser Leu Val Ile Thr
225                 230                 235                 240

Ala Glu Tyr Asp Pro Leu Arg Asp Gln Gly Glu Thr Tyr Ser His Ser
                250                 255                 260

Leu Asn Glu Ala Gly Asn Val Ser Thr Leu Val Arg Tyr Gln Gly Met
            265                 270                 275

Ile His Gly Phe Leu Ser Phe Tyr Glu Trp Ile Thr Ala Gly Lys Leu
        280                 285                 290

Ala Ile His His Ile Ala Gly Val Leu Arg Ser Val Leu
    295                 300                 305
```

(2) INFORMATION FOR SEQ ID NO:56:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 326 AMINO ACIDS
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

```
Val Ala Phe Phe Asp Met Pro Leu Glu Glu Leu Lys Lys Tyr Arg Pro
1               5                   10                  15

Glu Arg Tyr Glu Glu Lys Asp Phe Asp Glu Phe Trp Arg Glu Thr Leu
            20                  25                  30

Lys Glu Ser Glu Gly Phe Pro Leu Asp Pro Val Phe Glu Lys Val Asp
        35                  40                  45

Phe His Leu Lys Thr Val Glu Thr Tyr Asp Val Thr Phe Ser Gly Tyr
    50                  55                  60

Arg Gly Gln Arg Ile Lys Gly Trp Leu Leu Val Pro Lys Leu Ala Glu
65                  70                  75                  80

Glu Lys Leu Pro Cys Val Val Gln Tyr Ile Gly Tyr Asn Gly Gly Arg
                85                  90                  95

Gly Phe Pro His Asp Trp Leu Phe Trp Pro Ser Met Gly Tyr Ile Cys
            100                 105                 110



Phe Val Met Asp Thr Arg Gly Gln Gly Ser Gly Trp Met Lys Gly Asp
        115                 120                 125

Thr Pro Asp Tyr Pro Glu Gly Pro Val Asp Pro Gln Tyr Pro Gly Phe
    130                 135                 140

Met Thr Arg Gly Ile Leu Asp Pro Gly Thr Tyr Tyr Tyr Arg Arg Val
145                 150                 155                 160

Phe Val Asp Ala Val Arg Ala Val Glu Ala Ala Ile Ser Phe Pro Arg
                165                 170                 175

Val Asp Ser Arg Lys Val Val Val Ala Gly Gly Ser Gln Gly Gly Gly
            180                 185                 190

Ile Pro Leu Ala Val Ser Ala Leu Ser Asn Arg Val Lys Ala Leu Leu
        195                 200                 205

Cys Asp Val Pro Phe Leu Cys His Phe Arg Arg Ala Val Gln Leu Val
    210                 215                 220

Asp Thr His Pro Tyr Val Glu Ile Thr Asn Phe Leu Lys Thr His Arg
225                 230                 235                 240

Asp Lys Glu Glu Ile Val Phe Arg Thr Leu Ser Tyr Phe Asp Gly Val
                245                 250                 255

Asn Phe Ala Ala Arg Ala Lys Val Pro Ala Leu Phe Ser Val Gly Leu
            260                 265                 270

Met Asp Thr Ile Cys Pro Pro Ser Thr Val Phe Ala Ala Tyr Asn His
        275                 280                 285

Tyr Ala Gly Pro Lys Glu Ile Arg Ile Tyr Pro Tyr Asn Asn His Glu
    290                 295                 300

Gly Gly Gly Ser Phe Gln Ala Ile Glu Gln Val Lys Phe Leu Lys Arg
305                 310                 315                 320

Leu Phe Glu Glu Gly
                325
```

(2) INFORMATION FOR SEQ ID NO:57:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 298 AMINO ACIDS
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

```
Met Arg Thr Leu Ser Phe Gly Pro Met Thr Thr Gly Gly Ser Ile His
1               5                   10                  15

Met Ala Thr Met Asp Val Met Arg Gly Pro Gly Met Gln Arg Leu Ser
            20                  25                  30

Gln Gly Ala Arg Glu Ala Ala Asn His Pro Trp Ala Lys Arg Leu Gly
        35                  40                  45

Arg Met Gly Tyr Ala Ala Lys Gly Ala Val Tyr Ala Ile Ile Gly Val
    50                  55                  60

Leu Ala Leu Lys Leu Ala Ala Gly Glu Gly Gly Arg Thr Thr Asp Ser
65                  70                  75                  80

His Gly Ala Val Asn Thr Val Ala His Gly Pro Phe Gly Val Ala Leu
                85                  90                  95

Leu Ala Val Leu Val Val Gly Leu Leu Gly Tyr Val Val Trp Arg Phe
            100                 105                 110

Ala Gln Ala Phe Val Asp Thr Glu Asp Lys Gly Ser Asp Ala Lys Gly
        115                 120                 125


Ile Ala Thr Arg Ala Met Tyr Phe Leu Ser Gly Cys Ile Tyr Ala Ser
    130                 135                 140

Leu Ala Phe Phe Ala Ala Gln Ser Leu Val Gly Ala Ala His Gly Arg
145                 150                 155                 160

Ser Lys Gly Thr Gln Gly Trp Thr Ala Thr Leu Met Glu Gln Pro Phe
                165                 170                 175

Gly Arg Val Leu Val Ala Leu Val Gly Leu Gly Ile Val Gly Phe Ala
            180                 185                 190

Leu Lys Gln Phe His Thr Ala Trp Lys Ala Lys Phe Arg Glu Lys Leu
        195                 200                 205

Thr Leu Thr Gly Leu Ala Ala Arg Lys Gln His His Ile Glu Arg Met
    210                 215                 220

Cys Gln Phe Gly Ile Ala Ala Arg Gly Val Val Phe Ala Val Ile Gly
225                 230                 235                 240

Gly Phe Leu Val Arg Ser Ala Val Asp Ala Asn Pro Gly Glu Ala Lys
                245                 250                 255

Gly Leu Gly Glu Ala Leu Ala Val Val Ala Arg Gln Pro Ser Gly Asp
            260                 265                 270

Val Leu Leu Gly Val Val Ala Ala Gly Leu Val Ala Tyr Ala Ala Tyr
        275                 280                 285

Leu Phe Leu Gln Ala Arg Tyr Arg Glu Leu
    290                 295
```

(2) INFORMATION FOR SEQ ID NO:58:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 304 AMINO ACIDS
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

```
Met Ser Lys Phe Ala Ile Leu Trp Ala Leu Ile Thr Ala Tyr Leu Pro
1               5                   10                  15

Glu Pro Val Met Lys Leu Val Tyr Leu Gly Arg Arg Glu Thr Leu Gly
            20                  25                  30

Ala Arg Thr Leu Asp Val Lys Ala Gln Ala Val Gly Arg Leu Ala Asn
        35                  40                  45

Ala Thr Arg Pro Val Gly Val Ile Pro Thr Val Glu Glu Ser Arg Lys
    50                  55                  60

Met Thr Asp Lys Ala Val Ser Leu Phe Asp Gln Pro Ala Pro Glu Leu
65                  70                  75                  80

Phe Arg Lys Lys Asp Ile Gln Ile Asp Gly Ala Glu Gly Pro Ile Asp
                85                  90                  95

Ala Arg Ile Tyr Ser Gly Pro Ala Lys His Arg Pro Arg Pro Ile Leu
            100                 105                 110

Val Tyr Phe His Gly Gly Gly Trp Val Gln Gly Asn Leu Asp Ser His
        115                 120                 125

Asp Gly Val Cys Gly Lys Leu Ala Lys Trp Ala Asn Cys Ile Val Ile
    130                 135                 140

Ser Val Asp Tyr Arg Leu Ala Pro Glu His Lys Phe Pro Cys Ala Pro
145                 150                 155                 160

Leu Asp Ala Ile Ala Ala Tyr Lys Trp Val Arg Ala Asn Ala Thr Asn
                165                 170                 175
```

```
Leu Gly Gly Asp Pro Glu Arg Ile Gly Val Gly Gly Asp Ser Ala Gly
            180                 185                 190

Gly Asn Leu Ala Ala Val Val Cys Gln Gln Thr Ala Met Asn Gly Glu
        195                 200                 205

Arg Thr Pro Asp Leu Gln Val Leu Ile Tyr Pro Ala Leu Asp Ala Arg
    210                 215                 220

Met Ile Ser Thr Ser Met Glu Glu Leu Arg Asp Ala Tyr Ile Leu Pro
225                 230                 235                 240

Lys Ser Arg Met Glu Tyr Phe Leu Gly Leu Tyr Thr Arg Gly Pro Asp
                245                 250                 255

Asp Ile Glu Asp Leu Arg Met Ser Pro Ile Leu Arg Asp Thr Val Ala
            260                 265                 270

Asp Gln Pro Gln Ala Cys Ile Val Thr Cys Gly Phe Asp Pro Ala Arg
        275                 280                 285

Arg Arg Glu His Leu Arg Arg Thr Leu Asn Cys Arg Gly Asp Arg Arg
    290                 295                 300
```

(2) INFORMATION FOR SEQ ID NO:59:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 308 AMINO ACIDS
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

Val Ser Ile Arg Leu Arg Leu Leu Asn Trp Phe Leu Asn Thr Phe Glu
1 5 10 15

Lys Pro Lys Leu Ala Ala Ala Lys Thr Pro Asp Asp Leu Arg Lys Ser
20 25 30

Phe Glu Leu Lys Ala Arg Phe Leu Phe Pro Ala Pro Arg Lys Thr Arg
35 40 45

Phe Ser His Asp Val Leu Gln Ser Gly Ile Gly Ser Val Asn Ala Gln
50 55 60

Trp Ala Lys Ser Lys Ser Ala Ser Asp Asp Arg Val Ile Leu Tyr Phe
65 70 75 80

His Gly Gly Gly Tyr Val Phe Gly Ser Pro Lys Thr His Arg Ala Met
85 90 95

Leu Ala Arg Leu Ser Ala Met Thr Gly Leu Ser Ala Cys Leu Pro Asp
100 105 110

Tyr Arg Leu Ala Pro Glu His Pro Phe Pro Ala Ala Ile Glu Asp Ala
115 120 125

Val Leu Ser Tyr Lys Cys Leu Leu Glu Arg Ala Ile Glu Pro Gln Asn
130 135 140

Ile Ile Leu Gly Gly Asp Ser Ala Gly Gly Gly Leu Val Leu Ala Leu
145 150 155 160

Leu Ala Glu Ile Lys Ala Gln Ser Leu Pro Lys Pro Ala Gly Val Phe
165 170 175

Ala Leu Ser Pro Leu Val Asp Leu Ser Phe Ser Gly Leu Ser Phe Ser
180 185 190

Lys Asn Ala Gln Thr Asp Val Met Leu Pro Ala Ser Arg Ala Ala Asp
195 200 205

Met Ala Thr Leu Tyr Leu Asp Gly Ala Asp Ala Asp Asp Pro Arg Ala
210 215 220

Ser Pro Leu Gln Ala Asp Phe Ser Gly Met Pro Pro Val Phe Leu Thr
225 230 235 240

Ala Ser Asp Ser Glu Ile Leu Leu Asp Asp Cys Leu Arg Met Ala Asp
245 250 255

His Leu Arg Ala Gln Gly Val Val Val Thr Asp Arg Ile Val Glu Asn
260 265 270

His Pro His Val Trp His Ile Phe Gln Arg Leu Leu Pro Glu Ala Asp
275 280 285

Gln Gly Leu Arg Ala Ile Ala Ala Trp Ile Lys Pro Leu Leu Ser Gly
290 295 300

Ser Asn Glu Ser
305

(2) INFORMATION FOR SEQ ID NO:60:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 237 AMINO ACIDS
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

```
Met Leu Thr Phe Asn Val Leu Tyr Gly Met Met Lys Gln Lys Leu Ala
1               5                   10                  15

Ala Ile Leu Met Phe Leu Gly Leu Ser Ala Ala Glu Ala Gln Asp Trp
            20                  25                  30

Pro Asp Leu Gln Lys Tyr Arg Ser Ala Asn Lys Glu Ala Lys Leu Leu
        35                  40                  45

Pro Lys Glu Asn Arg Lys Val Val Phe Met Gly Asn Ser Ile Thr Glu
    50                  55                  60

Ala Trp Ile Ser Gln Arg Pro Glu Phe Phe Ser Glu Asn Gly Phe Ile
65                  70                  75                  80

Gly Arg Gly Ile Ser Gly Gln Thr Thr Pro Gln Met Leu Leu Arg Phe
                85                  90                  95

Arg Gln Asp Val Ile Asp Leu Gln Pro Lys Ala Val Val Ile Leu Ala
            100                 105                 110

Gly Thr Asn Asp Val Ala Gln Asn Thr Gly Pro Met Thr Ile Glu Glu
        115                 120                 125

Ser Leu Ala Asn Ile Lys Ser Met Val Glu Leu Ala Gln Ala Asn Gly
    130                 135                 140

Ile Thr Pro Val Leu Cys Thr Val Leu Pro Ala Asp Arg Phe Ser Trp
145                 150                 155                 160

Arg Pro Glu Leu Thr Pro Ala Glu Thr Ile Ile Ala Leu Asn Gln Leu
                165                 170                 175

Ile Lys Gln Tyr Ala Glu Ala Gln Gly Leu Ala Leu Val Asp Tyr His
            180                 185                 190

Ala Ala Leu Thr Asn Lys Gly Gly Gly Leu Pro Val Lys Tyr Gly Glu
        195                 200                 205

Asp Gly Val His Pro Asn Val Ala Gly Tyr Gln Val Met Glu Asn Ile
    210                 215                 220

Val Leu Pro Val Ile Ser Ser Glu Leu Ala Lys Leu Lys
225                 230                 235
```

(2) INFORMATION FOR SEQ ID NO:61:

(2) INFORMATION FOR SEQ ID NO:61:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 326 AMINO ACIDS
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

```
Met Ala Phe Phe Asp Leu Pro Leu Glu Glu Leu Lys Lys Tyr Arg Pro
1               5                   10                  15

Glu Arg Tyr Glu Glu Lys Asp Phe Asp Glu Phe Trp Glu Glu Thr Leu
            20                  25                  30

Ala Glu Ser Glu Lys Phe Pro Leu Asp Pro Val Phe Glu Arg Met Glu
        35                  40                  45

Ser His Leu Lys Thr Val Glu Ala Tyr Asp Val Thr Phe Ser Gly Tyr
    50                  55                  60

Arg Gly Gln Arg Ile Lys Gly Trp Leu Leu Val Pro Lys Leu Glu Glu
65                  70                  75                  80

Glu Lys Leu Pro Cys Val Val Gln Tyr Ile Gly Tyr Asn Gly Gly Arg
                85                  90                  95

Gly Phe Pro His Asp Trp Leu Phe Trp Pro Ser Met Gly Tyr Ile Cys
            100                 105                 110

Phe Val Met Asp Thr Arg Gly Gln Gly Ser Gly Trp Leu Lys Gly Asp
        115                 120                 125

Thr Pro Asp Tyr Pro Glu Gly Pro Val Asp Pro Gln Tyr Pro Gly Phe
    130                 135                 140

Met Thr Arg Gly Ile Leu Asp Pro Arg Thr Tyr Tyr Tyr Arg Arg Val
145                 150                 155                 160

Phe Thr Asp Ala Val Arg Ala Val Glu Ala Ala Ala Ser Phe Pro Gln
                165                 170                 175

Val Asp Gln Glu Arg Ile Val Ile Ala Gly Gly Ser Gln Gly Gly Gly
            180                 185                 190

Ile Ala Leu Ala Val Ser Ala Leu Ser Lys Lys Ala Lys Ala Leu Leu
        195                 200                 205

Cys Asp Val Pro Phe Leu Cys His Phe Arg Arg Ala Val Gln Leu Val
    210                 215                 220

Asp Thr His Pro Tyr Ala Glu Ile Thr Asn Phe Leu Lys Thr His Arg
225                 230                 235                 240

Asp Lys Glu Glu Ile Val Phe Arg Thr Leu Ser Tyr Phe Asp Gly Val
                245                 250                 255

Asn Phe Ala Ala Arg Ala Lys Ile Pro Ala Leu Phe Ser Val Gly Leu
            260                 265                 270

Met Asp Asn Ile Cys Pro Pro Ser Thr Val Phe Ala Ala Tyr Asn Tyr
        275                 280                 285

Tyr Ala Gly Pro Lys Glu Ile Arg Ile Tyr Pro Tyr Asn Asn His Glu
    290                 295                 300

Gly Gly Gly Ser Phe Gln Ala Val Glu Gln Val Lys Phe Leu Lys Lys
305                 310                 315                 320

Leu Phe Glu Lys Gly
                325
```

(2) INFORMATION FOR SEQ ID NO:62:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 220 AMINO ACIDS

(B) TYPE: AMINO ACID

(D) TOPOLOGY: LINEAR

```
(ii)  MOLECULE TYPE:  PROTEIN

(xi) SEQUENCE DESCRIPTION:  SEQ ID NO:62:

Leu Lys Tyr Phe Lys Ala Arg Leu Ala Gly Ile Thr Leu Leu Gly Leu
 1               5                  10                  15

Leu Ala Cys Thr Ser Ala Ser Ala Gln Thr Glu Pro Ile Val Phe Val
            20                  25                  30

His Gly Tyr Ser Gly Ser Ala Ser Asn Trp Asp Thr Met Leu Gly Arg
        35                  40                  45

Phe Arg Ser Asn Gly Tyr Ala Ser Gly Ser Leu Tyr Thr Phe Asn Tyr
    50                  55                  60

Asn Ser Leu Val Ser Ser Asn Arg Thr Ser Ala Ser Glu Leu Arg Ser
65                  70                  75                  80

Phe Val Asn Thr Val Arg Ser Arg His Gly Asn Ala Arg Ile Ala Leu
                85                  90                  95

Val Ala His Ser Asn Gly Gly Leu Val Ser Arg Trp Tyr Arg Ala Glu
            100                 105                 110

Leu Gly Gly Glu Thr Ala Thr Arg Arg Phe Val Thr Leu Gly Thr Pro
        115                 120                 125

His Arg Gly Thr Thr Trp Ala Tyr Ala Cys Tyr Ser Pro Ala Cys Phe
    130                 135                 140

Glu Met Arg Pro Gly Ser Ser Leu Leu Thr Thr Leu Gly Ser Arg Ala
145                 150                 155                 160

Cys Asp Arg Ser Leu Trp Ser Asn Thr Asp Gly Ile Ile Leu Pro Ala
                165                 170                 175

Ser Ser Ala Gln Cys Gly Val Ser Thr Arg Thr Ala Asp Val Ser His
            180                 185                 190

Leu Asp Leu Leu Thr Asp Ser Arg Val Tyr Thr Gln Leu Arg Thr Gln
        195                 200                 205

Leu Gln End Gly End Arg Cys Thr Glu Arg Ala Pro G
    210                 215                 220
```

SEQUENCE LISTING

[0133]

(1) GENERAL INFORMATION:

(i) APPLICANT: Diversa Corporation

(ii) TITLE OF INVENTION: ESTERASES

(iii) NUMBER OF SEQUENCES: 62

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Morrison & Foerster LLP
(B) STREET: 3811 Valley Centre Drive, Suite 500
(C) CITY: San Diego
(D) STATE: CA
(E) COUNTRY: USA
(F) ZIP: 92130

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Diskette
(B) COMPUTER: IBM Compatible
(C) OPERATING SYSTEM: Windows 95
(D) SOFTWARE: FastSEQ for Windows Version 2.0b

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: US97/02039
(B) FILING DATE: 11-FEB-1997

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: 08/602,359
(B) FILING DATE: 16-FEB-1996

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Einhorn, Ph.D., D.D.S., Gregory
(B) REGISTRATION NUMBER: 38,440
(C) REFERENCE/DOCKET NUMBER: 09010/010WO1

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 858 720 5133
(B) TELEFAX: 858 720 5125

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 52 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
CCGAGAATTC ATTAAAGAGG AGAAATTAAC TATGTCTTTA AACAAGCACT CT          52
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 31 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
CGGAAGATCT CTATCGTTTA GTGTATGATT T                                 31
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 52 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
CCGAGAATTC ATTAAAGAGG AGAAATTAAC TATGAAACTC CTTGAGCCCA CA          52
```

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 31 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```
CGGAAGATCT CGCCGGTACA CCATCAGCCA C                                 31
```

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 52 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

```
CCGAGAATTC ATTAAAGAGG AGAAATTAAC TATGCCATAT GTTAGGAATG GT          52
```

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 53 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

```
CGGAGGTACC TTAGAACTGT GCTGAAGAAA TAAATTCGTC CATTGCTCTA TTA              53
```

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 49 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

```
CGGAGGTACC TTAGAACTGT GCTGAAGAAA TAAATTCGTC CATTGCTCT                  49
```

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 53 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

```
CCGAGAATTC ATTAAAGAGG AGAAATTAAC TATGAGATTG AGGAAATTTG AAG              53
```

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 31 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

```
CGGAGGTACC CTATTCAGAA AGTACCTCTA A                                     31
```

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 52 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

```
CCGAGAATTC ATTAAAGAGG AGAAATTAAC TATGTTTAAT ATCAATGTCT TT                52
```

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 31 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

```
CGGAAGATCT TTAAGGATTT TCCCTGGGTA G                                       31
```

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 52 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

```
CCGAGAATTC ATTAAAGAGG AGAAATTAAC TATGGAGGTT TACAAGGCCA AA                52
```

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 31 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CGGAGGTACC TTATTGAGCC GAAGAGTACG A 31

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 53 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

CCGAGAATTC ATTAAAGAGG AGAAATTAAC TATGATTGGC AATTTGAAAT TGA 53

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 31 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

CGGAGGTACC TTAAAGTGCT CTCATATCCC C 31

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 52 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

CCGAGAATTC ATTAAAGAGG AGAAATTAAC TATGCCAGCT AATGACTCAC CC 52

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

CGGAAGATCT TCAACAGGCT CCAAATAATT TC 32

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CGGAAGATCT ACAGGCTCCA AATAATTTC 29

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 52 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

CCGAGAATTC ATTAAAGAGG AGAAATTAAC TATGCTTGAT ATGCCAATCG AC 52

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 31 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

CGGAGGTACC CTAGTCGAAC AGAAGAAGAG C 31

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 52 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

CCGAGAATTC ATTAAAGAGG AGAAATTAAC TATGCCCCTA GATCCTAGAA TT 52

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 31 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

CGGAGGTACC TTAAATTTTA TCATAAAATA C 31

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 555 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:

(A) NAME/KEY: Coding Sequence
(B) LOCATION: 1...552

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

```
ATG TCT TTA AAC AAG CAC TCT TGG ATG GAT ATG ATA ATA TTT ATT CTC       48
Met Ser Leu Asn Lys His Ser Trp Met Asp Met Ile Ile Phe Ile Leu
 1               5                  10                   15

AGC TTT TCT TTC CCA TTA ACA ATG ATC GCA TTA GCT ATC TCT ATG TCG       96
Ser Phe Ser Phe Pro Leu Thr Met Ile Ala Leu Ala Ile Ser Met Ser
             20                  25                  30

TCA TGG TTT AAT ATA TGG AAT AAT GCA TTA AGC GAT CTA GGA CAT GCT      144
Ser Trp Phe Asn Ile Trp Asn Asn Ala Leu Ser Asp Leu Gly His Ala
         35                  40                  45

GTT AAA AGC AGT GTT GCT CCA ATA TTC AAT CTA GGT CTT GCA ATT GGT      192
Val Lys Ser Ser Val Ala Pro Ile Phe Asn Leu Gly Leu Ala Ile Gly
     50                  55                  60

GGG ATA CTA ATT GTT ATA GTT GGT TTA AGA AAT CTT TAT TCG TGG AGT      240
Gly Ile Leu Ile Val Ile Val Gly Leu Arg Asn Leu Tyr Ser Trp Ser
 65                  70                  75                  80

AGA GTT AAA GGA TCT TTA ATC ATA TCC ATG GGT GTA TTT CTT AAC TTA      288
Arg Val Lys Gly Ser Leu Ile Ile Ser Met Gly Val Phe Leu Asn Leu
                 85                  90                  95

ATA GGG GTT TTC GAC GAA GTA TAT GGT TGG ATA CAT TTC CTA GTC TCA      336
Ile Gly Val Phe Asp Glu Val Tyr Gly Trp Ile His Phe Leu Val Ser
            100                 105                 110

GTA TTG TTT TTC TTA TCA ATA ATA GCA TAT TTC ATA GCT ATA TCA ATA      384
Val Leu Phe Phe Leu Ser Ile Ile Ala Tyr Phe Ile Ala Ile Ser Ile
        115                 120                 125

CTT GAC AAA TCA TGG ATA GCT GTT CTA CTA ATA ATA GGT CAT ATT GCA      432
Leu Asp Lys Ser Trp Ile Ala Val Leu Leu Ile Ile Gly His Ile Ala
    130                 135                 140

ATG TGG TAT CTA CAC TTT GCT TCA GAG ATT CCG AGA GGT GCG GCT ATT      480
Met Trp Tyr Leu His Phe Ala Ser Glu Ile Pro Arg Gly Ala Ala Ile
145                 150                 155                 160

CCC GAG TTA TTA GCG GTA TTC TCG TTT TTA CCA TTC TAT ATA AGA GAC      528
Pro Glu Leu Leu Ala Val Phe Ser Phe Leu Pro Phe Tyr Ile Arg Asp
                165                 170                 175

TAT TTT AAA TCA TAC ACT AAA CGA TAG                                  555
Tyr Phe Lys Ser Tyr Thr Lys Arg
            180
```

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1041 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:

(A) NAME/KEY: Coding Sequence
(B) LOCATION: 1...1038

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

```
ATG AAA CTC CTT GAG CCC ACA AAT ACC TCC TAC ACG CTG TTA CAG GAT        48
Met Lys Leu Leu Glu Pro Thr Asn Thr Ser Tyr Thr Leu Leu Gln Asp
 1               5                  10                  15

TTA GCA TTG CAT TTT GCA TTT TAC TGG TTT CTG GCC GTG TAT ACG TGG        96
```

```
Leu Ala Leu His Phe Ala Phe Tyr Trp Phe Leu Ala Val Tyr Thr Trp
            20                  25                  30

TTA CCC GGT GTC CTA GTC CGG GGC GTA GCT GTG GAC ACA GGG GTG GCT      144
Leu Pro Gly Val Leu Val Arg Gly Val Ala Val Asp Thr Gly Val Ala
        35                  40                  45

CGG GTG CCT GGG CTC GGC CGG CGC GGT AAG AGG CTG CTC CTG GCC GCT      192
Arg Val Pro Gly Leu Gly Arg Arg Gly Lys Arg Leu Leu Leu Ala Ala
    50                  55                  60

GTG GCT GTC TTG GCG CTT GTT GTG TCC GTT GTT GTC CCG GCT TAT GTG      240
Val Ala Val Leu Ala Leu Val Val Ser Val Val Val Pro Ala Tyr Val
65                  70                  75                  80

GCG TAT AGT AGT CTG CAC CCG GAG AGC TGT CGG CCC GTT GCG CCG GAG      288
Ala Tyr Ser Ser Leu His Pro Glu Ser Cys Arg Pro Val Ala Pro Glu
                85                  90                  95

GGG CTC ACC TAC AAA GAG TTC AGC GTG ACC GCG GAG GAT GGC TTG GTG      336
Gly Leu Thr Tyr Lys Glu Phe Ser Val Thr Ala Glu Asp Gly Leu Val
            100                 105                 110

GTT CGG GGC TGG GTG CTG GGC CCC GGC GCT GGG GGC AAC CCG GTG TTC      384
Val Arg Gly Trp Val Leu Gly Pro Gly Ala Gly Gly Asn Pro Val Phe
        115                 120                 125

GTT TTG ATG CAC GGG TAT ACT GGG TGC CGC TCG GCG CCC TAC ATG GCT      432
Val Leu Met His Gly Tyr Thr Gly Cys Arg Ser Ala Pro Tyr Met Ala
    130                 135                 140

GTG CTG GCC CGG GAG CTC GTG GAG TGG GGG TAC CCG GTG GTT GTG TTC      480
Val Leu Ala Arg Glu Leu Val Glu Trp Gly Tyr Pro Val Val Val Phe
145                 150                 155                 160

GAC TTC CGG GGC CAC GGG GAG AGC GGG GGC TCG ACG ACG ATT GGG CCC      528
Asp Phe Arg Gly His Gly Glu Ser Gly Gly Ser Thr Thr Ile Gly Pro
                165                 170                 175

CGG GAG GTG CTG GAT GCC CGG GCT GTG GTG GGC TAT GTC TCG GAG CGG      576
Arg Glu Val Leu Asp Ala Arg Ala Val Val Gly Tyr Val Ser Glu Arg
            180                 185                 190

TTC CCC GGC CGC CGG ATA ATA TTG GTG GGG TTC AGT ATG GGC GGC GCT      624
Phe Pro Gly Arg Arg Ile Ile Leu Val Gly Phe Ser Met Gly Gly Ala
        195                 200                 205

GTA GCG ATC GTG GAG GGT GCT GGG GAC CCG CGG GTC TAC GCG GTG GCT      672
Val Ala Ile Val Glu Gly Ala Gly Asp Pro Arg Val Tyr Ala Val Ala
    210                 215                 220

GCT GAT AGC CCG TAC TAT AGG CTC CGG GAC GTC ATA CCC CGG TGG CTG      720
Ala Asp Ser Pro Tyr Tyr Arg Leu Arg Asp Val Ile Pro Arg Trp Leu
225                 230                 235                 240
```

```
GAG TAC AAG ACG CCG CTG CCG GGC TGG GTG GGT GTG CTG GCC GGG TTC      768
Glu Tyr Lys Thr Pro Leu Pro Gly Trp Val Gly Val Leu Ala Gly Phe
                245                 250                 255

TAC GGG AGG CTG ATG GCG GGC GTT GAC CTC GGC TTC GGC CCC GCT GGG      816
Tyr Gly Arg Leu Met Ala Gly Val Asp Leu Gly Phe Gly Pro Ala Gly
            260                 265                 270

GTG GAG CGC GTG GAT AAG CCG TTG CTG GTG GTG TAT GGG CCC CGG GAC      864
Val Glu Arg Val Asp Lys Pro Leu Leu Val Val Tyr Gly Pro Arg Asp
        275                 280                 285

CCG CTG GTG ACG CGG GAC GAG GCG AGG AGC CTG GCG TCC CGT AGC CCG      912
Pro Leu Val Thr Arg Asp Glu Ala Arg Ser Leu Ala Ser Arg Ser Pro
    290                 295                 300

TGT GGC CGT CTC GTC GAG GTT CCT GGG GCT GGC CAC GTG GAG GCC GTG      960
Cys Gly Arg Leu Val Glu Val Pro Gly Ala Gly His Val Glu Ala Val
305                 310                 315                 320

GAT GTG CTC GGG CCG GGC CGC TAC GCA GAC ATG CTG ATA GAG CTG GCG     1008
Asp Val Leu Gly Pro Gly Arg Tyr Ala Asp Met Leu Ile Glu Leu Ala
                325                 330                 335

CAC GAG GAG TGC CCT CCG GGG GCC GGT GGC TGA                         1041
His Glu Glu Cys Pro Pro Gly Ala Gly Gly
            340                 345
```

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 789 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:

(A) NAME/KEY: Coding Sequence
(B) LOCATION: 1...786

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

```
ATG CCA TAT GTT AGG AAT GGT GGT GTA AAT ATC TAT TAT GAA CTG GTG       48
Met Pro Tyr Val Arg Asn Gly Gly Val Asn Ile Tyr Tyr Glu Leu Val
 1               5                  10                  15

GAT GGA CCT GAG CCA CCA ATT GTC TTT GTT CAC GGA TGG ACA GCA AAT       96
Asp Gly Pro Glu Pro Pro Ile Val Phe Val His Gly Trp Thr Ala Asn
            20                  25                  30

ATG AAT TTT TGG AAA GAG CAA AGA CGT TAT TTT GCA GGC AGG AAT ATG      144
Met Asn Phe Trp Lys Glu Gln Arg Arg Tyr Phe Ala Gly Arg Asn Met
```

```
         35                  40                  45

ATG TTG TTT GTC GAT AAC AGA GGT CAT GGC AGG TCC GAT AAG CCA CTT      192
Met Leu Phe Val Asp Asn Arg Gly His Gly Arg Ser Asp Lys Pro Leu
     50                  55                  60

GGA TAC GAT TTC TAC AGA TTT GAG AAC TTC ATT TCA GAT TTA GAT GCG      240
Gly Tyr Asp Phe Tyr Arg Phe Glu Asn Phe Ile Ser Asp Leu Asp Ala
 65                  70                  75                  80

GTT GTT AGG GAG ACT GGA GTG GAG AAA TTT GTT CTC GTC GGA CAT TCA      288
Val Val Arg Glu Thr Gly Val Glu Lys Phe Val Leu Val Gly His Ser
                 85                  90                  95

TTC GGA ACA ATG ATC TCT ATG AAG TAC TGT TCG GAG TAT CGG AAT CGG      336
Phe Gly Thr Met Ile Ser Met Lys Tyr Cys Ser Glu Tyr Arg Asn Arg
            100                 105                 110

GTT CTT GCT CTA ATC CTC ATA GGT GGT GGG AGC AGA ATA AAG CTT CTA      384
Val Leu Ala Leu Ile Leu Ile Gly Gly Gly Ser Arg Ile Lys Leu Leu
        115                 120                 125

CAC AGA ATT GGA TAT CCT TTA GCA AAG ATT CTT GCA TCC ATT GCA TAC      432
His Arg Ile Gly Tyr Pro Leu Ala Lys Ile Leu Ala Ser Ile Ala Tyr
    130                 135                 140

AAG AAG TCT TCA AGA TTG GTC GCA GAT CTT TCC TTT GGC AAA AAT GCT      480
Lys Lys Ser Ser Arg Leu Val Ala Asp Leu Ser Phe Gly Lys Asn Ala
145                 150                 155                 160

GGT GAA CTT AAA GAG TGG GGA TGG AAA CAG GCA ATG GAT TAT ACA CCC      528
Gly Glu Leu Lys Glu Trp Gly Trp Lys Gln Ala Met Asp Tyr Thr Pro
                165                 170                 175

TCC TAC GTG GCA ATG TAC ACG TAC AGA ACT CTA ACG AAA GTG AAT CTT      576
Ser Tyr Val Ala Met Tyr Thr Tyr Arg Thr Leu Thr Lys Val Asn Leu
            180                 185                 190

GAA AAT ATC TTG GAG AAA ATA GAC TGT CCA ACA CTG ATT ATC GTT GGA      624
Glu Asn Ile Leu Glu Lys Ile Asp Cys Pro Thr Leu Ile Ile Val Gly
        195                 200                 205

GAA GAG GAT GCA CTA TTG CCC GTT AGC AAA TCA GTT GAG CTG AGC AGG      672
Glu Glu Asp Ala Leu Leu Pro Val Ser Lys Ser Val Glu Leu Ser Arg
    210                 215                 220

AGG ATA GAA AAC TCA AAG CTT GTG ATC ATC CCA AAC TCG GGG CAT TGC      720
Arg Ile Glu Asn Ser Lys Leu Val Ile Ile Pro Asn Ser Gly His Cys
225                 230                 235                 240

GTA ATG CTT GAG AGT CCA AGT GAG GTT AAT AGA GCA ATG GAC GAA TTC      768
Val Met Leu Glu Ser Pro Ser Glu Val Asn Arg Ala Met Asp Glu Phe
                245                 250                 255

ATT TCT TCA GCA CAG TTC TAA                                          789
```

```
Ile Ser Ser Ala Gln Phe
            260
```

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 756 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:

(A) NAME/KEY: Coding Sequence
(B) LOCATION: 1...753

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

```
TTG AGA TTG AGG AAA TTT GAA GAG ATA AAC CTC GTT CTT TCG GGA GGA       48
Leu Arg Leu Arg Lys Phe Glu Glu Ile Asn Leu Val Leu Ser Gly Gly
 1               5                  10                  15

GCT GCA AAG GGC ATA GCC CAC ATA GGT GTT TTG AAA GCT ATA AAC GAG       96
Ala Ala Lys Gly Ile Ala His Ile Gly Val Leu Lys Ala Ile Asn Glu
            20                  25                  30

CTC GGT ATA AGG GTG AGG GCT TTA AGC GGG GTG AGC GCC GGG GCA ATC      144
Leu Gly Ile Arg Val Arg Ala Leu Ser Gly Val Ser Ala Gly Ala Ile
        35                  40                  45

GTT TCG GTC TTT TAT GCC TCA GGC TAC TCC CCT GAA GGG ATG TTC AGC      192
Val Ser Val Phe Tyr Ala Ser Gly Tyr Ser Pro Glu Gly Met Phe Ser
    50                  55                  60

CTT CTG AAG AGG GTA AAC TGG CTG AAG CTG TTT AAG TTC AAG CCA CCT      240
Leu Leu Lys Arg Val Asn Trp Leu Lys Leu Phe Lys Phe Lys Pro Pro
65                  70                  75                  80

CTG AAG GGA TTG ATA GGG TGG GAG AAG GCT ATA AGA TTC CTT GAG GAA      288
Leu Lys Gly Leu Ile Gly Trp Glu Lys Ala Ile Arg Phe Leu Glu Glu
                85                  90                  95

GTT CTC CCT TAC AGG AGA ATA GAA AAA CTT GAG ATA CCG ACG TAT ATA      336
Val Leu Pro Tyr Arg Arg Ile Glu Lys Leu Glu Ile Pro Thr Tyr Ile
            100                 105                 110

TGC GCG ACG GAT TTA TAC TCG GGA AGG GCT CTA TAC CTC TCG GAA GGG      384
Cys Ala Thr Asp Leu Tyr Ser Gly Arg Ala Leu Tyr Leu Ser Glu Gly
        115                 120                 125

AGT TTA ATC CCC GCA CTT CTC GGC AGC TGT GCA ATT CCC GGC ATA TTT      432
Ser Leu Ile Pro Ala Leu Leu Gly Ser Cys Ala Ile Pro Gly Ile Phe
    130                 135                 140
```

```
GAA CCC GTT GAG TAT AAG AAT TAC TTG CTC GTT GAC GGA GGT ATA GTT       480
Glu Pro Val Glu Tyr Lys Asn Tyr Leu Leu Val Asp Gly Gly Ile Val
145                 150                 155                 160

AAC AAC CTT CCC GTT GAG CCC TTT CAG GAA AGC GGT ATT CCC ACC GTT       528
Asn Asn Leu Pro Val Glu Pro Phe Gln Glu Ser Gly Ile Pro Thr Val
                165                 170                 175

TGC GTT GAT GTC CTT CCC ATA GAG CCG GAA AAG GAT ATA AAG AAC ATT       576
Cys Val Asp Val Leu Pro Ile Glu Pro Glu Lys Asp Ile Lys Asn Ile
            180                 185                 190

CTT CAC ATC CTT TTG AGG AGC TTC TTT CTT GCG GTC CGC TCA AAC TCC       624
Leu His Ile Leu Leu Arg Ser Phe Phe Leu Ala Val Arg Ser Asn Ser
        195                 200                 205

GAA AAG AGA AAG GAG TTT TGT GAC CTC GTT ATA GTT CCT GAG CTT GAG       672
Glu Lys Arg Lys Glu Phe Cys Asp Leu Val Ile Val Pro Glu Leu Glu
    210                 215                 220

GAG TTC ACA CCC CTT GAT GTT AGA AAA GCG GAC CAA ATA ATG GAG AGG       720
Glu Phe Thr Pro Leu Asp Val Arg Lys Ala Asp Gln Ile Met Glu Arg
225                 230                 235                 240

GGA TAC ATA AAG GCC TTA GAG GTA CTT TCT GAA TAG                       756
Gly Tyr Ile Lys Ala Leu Glu Val Leu Ser Glu
                245                 250
```

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 894 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:

(A) NAME/KEY: Coding Sequence
(B) LOCATION: 1...891

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

```
ATG TTT AAT ATC AAT GTC TTT GTT AAT ATA TCT TGG CTG TAT TTT TCA       48
Met Phe Asn Ile Asn Val Phe Val Asn Ile Ser Trp Leu Tyr Phe Ser
 1               5                  10                  15

GGG ATA GTT ATG AAG ACT GTG GAA GAG TAT GCG CTA CTT GAA ACA GGC       96
Gly Ile Val Met Lys Thr Val Glu Glu Tyr Ala Leu Leu Glu Thr Gly
            20                  25                  30
```

```
GTA AGA GTG TTT TAT CGG TGT GTA ATC CCG GAG AAA GCT TTT AAC ACT       144
Val Arg Val Phe Tyr Arg Cys Val Ile Pro Glu Lys Ala Phe Asn Thr
         35                  40                  45

TTG ATA ATA GGT TCA CAC GGA TTG GGG GCG CAC AGT GGA ATC TAC ATT       192
Leu Ile Ile Gly Ser His Gly Leu Gly Ala His Ser Gly Ile Tyr Ile
     50                  55                  60

AGT GTT GCT GAA GAA TTT GCT AGG CAC GGA TTT GGA TTC TGC ATG CAC       240
Ser Val Ala Glu Glu Phe Ala Arg His Gly Phe Gly Phe Cys Met His
 65                  70                  75                  80

GAT CAA AGG GGA CAT GGG AGA ACG GCA AGC GAT AGA GAA AGA GGG TAT       288
Asp Gln Arg Gly His Gly Arg Thr Ala Ser Asp Arg Glu Arg Gly Tyr
                 85                  90                  95

GTG GAG GGC TTT CAC AAC TTC ATA GAG GAT ATG AAG GCC TTC TCC GAT       336
Val Glu Gly Phe His Asn Phe Ile Glu Asp Met Lys Ala Phe Ser Asp
            100                 105                 110

TAT GCC AAG TGG CGC GTG GGA GGT GAC GAA ATA ATA TTG CTA GGA CAC       384
Tyr Ala Lys Trp Arg Val Gly Gly Asp Glu Ile Ile Leu Leu Gly His
        115                 120                 125

AGT ATG GGC GGG CTG ATA GCG CTC TTA ACA GTT GCA ACT TAT AAA GAA       432
Ser Met Gly Gly Leu Ile Ala Leu Leu Thr Val Ala Thr Tyr Lys Glu
    130                 135                 140

ATC GCC AAG GGA GTT ATC GCG CTA GCC CCG GCC CTC CAA ATC CCC TTA       480
Ile Ala Lys Gly Val Ile Ala Leu Ala Pro Ala Leu Gln Ile Pro Leu
145                 150                 155                 160

ACC CCG GCT AGA AGA CTT GTT CTA AGC CTC GCG TCA AGG CTT GCC CCG       528
Thr Pro Ala Arg Arg Leu Val Leu Ser Leu Ala Ser Arg Leu Ala Pro
                165                 170                 175

CAT TCT AAG ATC ACC TTA CAA AGG AGA TTG CCG CAG AAA CCA GAG GGT       576
His Ser Lys Ile Thr Leu Gln Arg Arg Leu Pro Gln Lys Pro Glu Gly
            180                 185                 190

TTT CAA AGA GCA AAA GAT ATA GAA TAC AGT CTG AGT GAA ATA TCA GTC       624
Phe Gln Arg Ala Lys Asp Ile Glu Tyr Ser Leu Ser Glu Ile Ser Val
        195                 200                 205

AAG CTC GTG GAC GAA ATG ATT AAA GCA TCA TCT ATG TTC TGG ACC ATA       672
Lys Leu Val Asp Glu Met Ile Lys Ala Ser Ser Met Phe Trp Thr Ile
    210                 215                 220

GCA GGG GAA ATT AAT ACT CCC GTC CTG CTT ATT CAT GGG GAA AAA GAC       720
Ala Gly Glu Ile Asn Thr Pro Val Leu Leu Ile His Gly Glu Lys Asp
225                 230                 235                 240

AAT GTC ATA CCT CCG GAG GCG AGC AAA AAA GCC TAC CAA TTA ATA CCT       768
Asn Val Ile Pro Pro Glu Ala Ser Lys Lys Ala Tyr Gln Leu Ile Pro
```

```
                245                 250                 255

TCA TTC CCT AAA GAG TTG AAA ATA TAC CCC GAT CTT GGA CAC AAC TTG        816
Ser Phe Pro Lys Glu Leu Lys Ile Tyr Pro Asp Leu Gly His Asn Leu
            260                 265                 270

TTT TTT GAA CCA GGC GCG GTG AAA ATC GTC ACA GAC ATT GTA GAG TGG        864
Phe Phe Glu Pro Gly Ala Val Lys Ile Val Thr Asp Ile Val Glu Trp
        275                 280                 285

GTT AAG AAT CTA CCC AGG GAA AAT CCT TAA                                894
Val Lys Asn Leu Pro Arg Glu Asn Pro
    290                 295
```

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 789 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA
(ix) FEATURE:

(A) NAME/KEY: Coding Sequence
(B) LOCATION: 1...786

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

```
ATG GAG GTT TAC AAG GCC AAA TTC GGC GAA GCA AAG CTC GGC TGG GTC         48
Met Glu Val Tyr Lys Ala Lys Phe Gly Glu Ala Lys Leu Gly Trp Val
 1               5                  10                  15

GTT CTG GTT CAT GGC CTC GGC GAG CAC AGC GGA AGG TAT GGA AGA CTG         96
Val Leu Val His Gly Leu Gly Glu His Ser Gly Arg Tyr Gly Arg Leu
            20                  25                  30
ATT AAG GAA CTC AAC TAT GCC GGC TTT GGA GTT TAC ACC TTC GAC TGG        144
Ile Lys Glu Leu Asn Tyr Ala Gly Phe Gly Val Tyr Thr Phe Asp Trp
        35                  40                  45

CCC GGC CAC GGG AAG AGC CCG GGC AAG AGA GGG CAC ACG AGC GTC GAG        192
Pro Gly His Gly Lys Ser Pro Gly Lys Arg Gly His Thr Ser Val Glu
    50                  55                  60

GAG GCG ATG GAA ATC ATC GAC TCG ATA ATC GAG GAG ATC AGG GAG AAG        240
Glu Ala Met Glu Ile Ile Asp Ser Ile Ile Glu Glu Ile Arg Glu Lys
65                  70                  75                  80

CCC TTC CTC TTC GGC CAC AGC CTC GGT GGT CTA ACT GTC ATC AGG TAC        288
Pro Phe Leu Phe Gly His Ser Leu Gly Gly Leu Thr Val Ile Arg Tyr
                85                  90                  95

GCT GAG ACG CGG CCC GAT AAA ATA CGG GGA TTA ATA GCT TCC TCG CCT        336
```

```
Ala Glu Thr Arg Pro Asp Lys Ile Arg Gly Leu Ile Ala Ser Ser Pro
            100                 105                 110

GCC CTC GCC AAG AGC CCG GAA ACG CCG GGC TTC ATG GTG GCC CTC GCG      384
Ala Leu Ala Lys Ser Pro Glu Thr Pro Gly Phe Met Val Ala Leu Ala
        115                 120                 125

AAG TTC CTT GGA AAG ATC GCC CCG GGA GTT GTT CTC TCC AAC GGC ATA      432
Lys Phe Leu Gly Lys Ile Ala Pro Gly Val Val Leu Ser Asn Gly Ile
    130                 135                 140

AAG CCG GAA CTC CTC TCG AGG AAC AGG GAC GCC GTG AGG AGG TAC GTT      480
Lys Pro Glu Leu Leu Ser Arg Asn Arg Asp Ala Val Arg Arg Tyr Val
145                 150                 155                 160

GAA GAC CCA CTC GTC CAC GAC AGG ATT TCG GCC AAG CTG GGA AGG AGC      528
Glu Asp Pro Leu Val His Asp Arg Ile Ser Ala Lys Leu Gly Arg Ser
                165                 170                 175

ATC TTC GTG AAC ATG GAG CTG GCC CAC AGG GAG GCG GAC AAG ATA AAA      576
Ile Phe Val Asn Met Glu Leu Ala His Arg Glu Ala Asp Lys Ile Lys
            180                 185                 190

GTC CCG ATC CTC CTT CTG ATC GGC ACT GGC GAT GTA ATA ACC CCG CCT      624
Val Pro Ile Leu Leu Leu Ile Gly Thr Gly Asp Val Ile Thr Pro Pro
        195                 200                 205

GAA GGC TCA CGC AGA CTC TTC GAG GAG CTG GCC GTC GAG AAC AAA ACC      672
Glu Gly Ser Arg Arg Leu Phe Glu Glu Leu Ala Val Glu Asn Lys Thr
    210                 215                 220

CTG AGG GAG TTC GAG GGG GCG TAC CAC GAG ATA TTT GAA GAC CCC GAG      720
Leu Arg Glu Phe Glu Gly Ala Tyr His Glu Ile Phe Glu Asp Pro Glu
225                 230                 235                 240

TGG GCC GAG GAG TTC CAC GAA ACA ATT GTT AAG TGG CTG GTT GAA AAA      768
Trp Ala Glu Glu Phe His Glu Thr Ile Val Lys Trp Leu Val Glu Lys
                245                 250                 255

TCG TAC TCT TCG GCT CAA TAA                                          789
Ser Tyr Ser Ser Ala Gln
            260
```

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 750 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:

(A) NAME/KEY: Coding Sequence
(B) LOCATION: 1...747

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

```
TTG ATT GGC AAT TTG AAA TTG AAG AGG TTT GAA GAG GTT AAC TTA GTT        48
Leu Ile Gly Asn Leu Lys Leu Lys Arg Phe Glu Glu Val Asn Leu Val
 1               5                   10                  15

CTT TCG GGA GGG GCT GCC AAG GGT ATC GCC CAT ATA GGT GTT TTA AAA        96
Leu Ser Gly Gly Ala Ala Lys Gly Ile Ala His Ile Gly Val Leu Lys
             20                  25                  30

GCT CTG GAA GAG CTC GGT ATA AAG GTA AAG AGG CTC AGC GGG GTA AGT       144
Ala Leu Glu Glu Leu Gly Ile Lys Val Lys Arg Leu Ser Gly Val Ser
         35                  40                  45

GCT GGA GCT ATC GTT TCC GTC TTT TAC GCT TCG GGC TAC ACT CCC GAC       192
Ala Gly Ala Ile Val Ser Val Phe Tyr Ala Ser Gly Tyr Thr Pro Asp
     50                  55                  60

GAG ATG TTA AAA CTC CTG AAA GAG GTA AAC TGG CTC AAA CTT TTT AAG       240
Glu Met Leu Lys Leu Leu Lys Glu Val Asn Trp Leu Lys Leu Phe Lys
 65                  70                  75                  80

TTC AAA ACA CCG AAA ATG GGC TTA ATG GGG TGG GAG AAG GCT GCA GAG       288
Phe Lys Thr Pro Lys Met Gly Leu Met Gly Trp Glu Lys Ala Ala Glu
                 85                  90                  95

TTT TTG GAA AAA GAG CTC GGA GTT AAG AGG CTG GAA GAC CTG AAC ATA       336
Phe Leu Glu Lys Glu Leu Gly Val Lys Arg Leu Glu Asp Leu Asn Ile
            100                 105                 110

CCA ACC TAT CTT TGC TCG GCG GAT CTG TAC ACG GGA AAG GCT CTT TAC       384
Pro Thr Tyr Leu Cys Ser Ala Asp Leu Tyr Thr Gly Lys Ala Leu Tyr
        115                 120                 125
TTC GGC AGA GGT GAC TTA ATT CCC GTG CTT CTC GGA AGT TGT TCC ATA       432
Phe Gly Arg Gly Asp Leu Ile Pro Val Leu Leu Gly Ser Cys Ser Ile
    130                 135                 140

CCC GGG ATT TTT GAA CCA GTT GAG TAC GAG AAT TTT CTA CTT GTT GAC       480
Pro Gly Ile Phe Glu Pro Val Glu Tyr Glu Asn Phe Leu Leu Val Asp
145                 150                 155                 160

GGA GGT ATA GTG AAC AAC CTG CCC GTA GAA CCT TTG GAA AAG TTC AAA       528
Gly Gly Ile Val Asn Asn Leu Pro Val Glu Pro Leu Glu Lys Phe Lys
                165                 170                 175

GAA CCC ATA ATC GGG GTA GAT GTG CTT CCC ATA ACT CAA GAA AGA AAG       576
Glu Pro Ile Ile Gly Val Asp Val Leu Pro Ile Thr Gln Glu Arg Lys
            180                 185                 190

ATT AAA AAT ATA CTC CAC ATC CTT ATA AGG AGC TTC TTT CTG GCG GTT       624
Ile Lys Asn Ile Leu His Ile Leu Ile Arg Ser Phe Phe Leu Ala Val
        195                 200                 205
```

```
CGT TCC AAT TCG GAA AAG AGA AAG GAG TTC TGC AAC GTA GTT ATA GAA       672
Arg Ser Asn Ser Glu Lys Arg Lys Glu Phe Cys Asn Val Val Ile Glu
    210                 215                     220

CCT CCC CTT GAA GAG TTC TCT CCT CTG GAC GTA AAT AAG GCG GAC GAG       720
Pro Pro Leu Glu Glu Phe Ser Pro Leu Asp Val Asn Lys Ala Asp Glu
225                 230                     235                 240

ATA TTC TGC GGG GAT ATG AGA GCA CTT TAA                               750
Ile Phe Cys Gly Asp Met Arg Ala Leu
                245
```

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1017 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:

(A) NAME/KEY: Coding Sequence
(B) LOCATION: 1...1014

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

```
ATG CCA GCT AAT GAC TCA CCC ACG ATC GAC TTT AAT CCT CGC GGC ATT        48
Met Pro Ala Asn Asp Ser Pro Thr Ile Asp Phe Asn Pro Arg Gly Ile
  1               5                  10                  15

CTT CGC AAC GCT CAC GCA CAG GTT ATT TTA GCG ACT TCC GGC TTG CGC        96
Leu Arg Asn Ala His Ala Gln Val Ile Leu Ala Thr Ser Gly Leu Arg
             20                  25                  30

AAA GCG TTT TTG AAA CGC ACG CAC AAG AGC TAC CTC AGC ACT GCC CAA       144
Lys Ala Phe Leu Lys Arg Thr His Lys Ser Tyr Leu Ser Thr Ala Gln
        35                  40                  45

TGG CTG GAG CTC GAT GCC GGC AAC GGA GTT ACC TTG GCC GGA GAG CTT       192
Trp Leu Glu Leu Asp Ala Gly Asn Gly Val Thr Leu Ala Gly Glu Leu
    50                  55                  60

AAC ACA GCG CCT GCA ACT GCA TCC TCC TCC CAC CCG GCG CAC AAG AAC       240
Asn Thr Ala Pro Ala Thr Ala Ser Ser Ser His Pro Ala His Lys Asn
 65                  70                  75                  80

ACT CTG GTT ATT GTG CTG CAC GGC TGG GAA GGC TCC AGC CAG TCG GCC       288
Thr Leu Val Ile Val Leu His Gly Trp Glu Gly Ser Ser Gln Ser Ala
                 85                  90                  95
```

```
TAT GCG ACC TCC GCT GGC AGC ACG CTT TTC GAC AAT GGG TTC GAC ACT        336
Tyr Ala Thr Ser Ala Gly Ser Thr Leu Phe Asp Asn Gly Phe Asp Thr
            100                 105                 110

TTT CGC CTT AAT TTT CGC GAT CAC GGC GAC ACC TAC CAC TTA AAC CGC        384
Phe Arg Leu Asn Phe Arg Asp His Gly Asp Thr Tyr His Leu Asn Arg
        115                 120                 125

GGC ATA TTT AAC TCA TCG CTG ATT GAC GAA GTA GTG GGC GCA GTC AAA        432
Gly Ile Phe Asn Ser Ser Leu Ile Asp Glu Val Val Gly Ala Val Lys
    130                 135                 140

GCC ATC CAG CAG CAA ACC GAC TAC GAC AAG TAT TGC CTG ATG GGG TTC        480
Ala Ile Gln Gln Gln Thr Asp Tyr Asp Lys Tyr Cys Leu Met Gly Phe
145                 150                 155                 160

TCA CTG GGT GGG AAC TTT GCC TTG CGC GTC GCG GTG CGG GAA CAG CAT        528
Ser Leu Gly Gly Asn Phe Ala Leu Arg Val Ala Val Arg Glu Gln His
                165                 170                 175

CTC GCT AAA CCG CTA GCG GGC GTG CTC GCC GTA TGC CCG GTA CTC GAC        576
Leu Ala Lys Pro Leu Ala Gly Val Leu Ala Val Cys Pro Val Leu Asp
            180                 185                 190

CCC GCA CAC ACC ATG ATG GCC CTA AAC CGA GGT GCG TTT TTC TAC GGC        624
Pro Ala His Thr Met Met Ala Leu Asn Arg Gly Ala Phe Phe Tyr Gly
        195                 200                 205

CGC TAT TTT GCG CAT AAA TGG AAG CGC TCG TTA ACC GCA AAA CTT GCA        672
Arg Tyr Phe Ala His Lys Trp Lys Arg Ser Leu Thr Ala Lys Leu Ala
    210                 215                 220

GCT TTC CCA GAC TAC AAA TAC GGC AAA GAT TTA AAA TCG ATA CAC ACG        720
Ala Phe Pro Asp Tyr Lys Tyr Gly Lys Asp Leu Lys Ser Ile His Thr
225                 230                 235                 240

CTT GAT GAG TTA AAC AAC TAT TTC ATT CCC CGC TAC ACC GGC TTC AAC        768
Leu Asp Glu Leu Asn Asn Tyr Phe Ile Pro Arg Tyr Thr Gly Phe Asn
                245                 250                 255

TCA GTC TCC GAA TAC TTC AAA AGT TAC ACG CTC ACC GGG CAG AAG CTC        816
Ser Val Ser Glu Tyr Phe Lys Ser Tyr Thr Leu Thr Gly Gln Lys Leu
            260                 265                 270

GCG TTT CTC AAC TGC CCC AGT TAC ATT CTG GCA GCT GGC GAC GAC CCA        864
Ala Phe Leu Asn Cys Pro Ser Tyr Ile Leu Ala Ala Gly Asp Asp Pro
        275                 280                 285

ATA ATT CCA GCA TCC GAC TTT CAG AAA ATA GCC AAG CCT GCG AAT CTG        912
Ile Ile Pro Ala Ser Asp Phe Gln Lys Ile Ala Lys Pro Ala Asn Leu
    290                 295                 300

CAC ATA ACA GTA ACG CAA CAA GGT TCT CAT TGC GCA TAC CTG GAA AAC        960
His Ile Thr Val Thr Gln Gln Gly Ser His Cys Ala Tyr Leu Glu Asn
305                 310                 315                 320
```

```
CTG CAT AAA CCT AGT GCT GCC GAC AAA TAT GCG GTG AAA TTA TTT GGA      1008
Leu His Lys Pro Ser Ala Ala Asp Lys Tyr Ala Val Lys Leu Phe Gly
                325                 330                 335

GCC TGT TGA                                                          1017
Ala Cys
```

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 936 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:

(A) NAME/KEY: Coding Sequence
(B) LOCATION: 1...933

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

```
ATG CTT GAT ATG CCA ATC GAC CCT GTT TAC TAC CAG CTT GCT GAG TAT        48
Met Leu Asp Met Pro Ile Asp Pro Val Tyr Tyr Gln Leu Ala Glu Tyr
 1               5                  10                  15

TTC GAC AGT CTG CCG AAG TTC GAC CAG TTT TCC TCG GCC AGA GAG TAC        96
Phe Asp Ser Leu Pro Lys Phe Asp Gln Phe Ser Ser Ala Arg Glu Tyr
            20                  25                  30
AGG GAG GCG ATA AAT CGA ATA TAC GAG GAG AGA AAC CGG CAG CTG AGC       144
Arg Glu Ala Ile Asn Arg Ile Tyr Glu Glu Arg Asn Arg Gln Leu Ser
        35                  40                  45

CAG CAT GAG AGG GTT GAA AGA GTT GAG GAC AGG ACG ATT AAG GGG AGG       192
Gln His Glu Arg Val Glu Arg Val Glu Asp Arg Thr Ile Lys Gly Arg
    50                  55                  60

AAC GGA GAC ATC AGA GTC AGA GTT TAC CAG CAG AAG CCC GAT TCC CCG       240
Asn Gly Asp Ile Arg Val Arg Val Tyr Gln Gln Lys Pro Asp Ser Pro
65                  70                  75                  80

GGT CTG GTT TAC TAT CAC GGT GGT GGA TTT GTG ATT TGC AGC ATC GAG       288
Val Leu Val Tyr Tyr His Gly Gly Gly Phe Val Ile Cys Ser Ile Glu
                85                  90                  95

TCG CAC GAC GCC TTA TGC AGG AGA AYY GCG AGA CTT TCA AAC TCT ACC       336
Ser His Asp Ala Leu Cys Arg Arg Ile Ala Arg Leu Ser Asn Ser Thr
            100                 105                 110

GTA GTC TCC GTG GAT TAC AGG CTC GCT CCT GAG CAC AAG TTT CCC CCC       384
Val Val Ser Val Asp Tyr Arg Leu Ala Pro Glu His Lys Phe Pro Ala
```

```
        115                 120                 125

CCA GTT TAT CAT TGC TAC GAT GCG ACC AAG TGG GTT GCT GAG AAC CGG        432
Ala Val Tyr Asp Cys Tyr Asp Ala Thr Lys Trp Val Ala Glu Asn Ala
    130                 135                 140

GAG GAG CTG AGG ATT GAC CCG TCA AAA ATC TTC GTT GGG GGG GAC AGT        480
Glu Glu Leu Arg Ile Asp Pro Ser Lys Ile Phe Val Gly Gly Asp Ser
145                 150                 155                 160

GCG GGA CG  AAT CTT GCC CGG GCG CTT TCA ATA ATG GCG AGA GAC AGC        528
Ala Gly Gly Asn Leu Ala Ala Ala Val Ser Ile Met Ala Arg Asp Ser
                165                 170                 175

GGA GAA GAT TTC ATA AAG CAT CAA ATT CTA ACT TAC CCC GTT GTG AAC        576
Gly Glu Asp Phe Ile Lys His Gln Ile Leu Ile Tyr Pro Val Val Asn
            180                 185                 190

TTT GTA GCC CCC ACA CCA TCG CTT CTG GAG TTT GGA GAG GGG CTG TGG        624
Phe Val Ala Pro Thr Pro Ser Leu Leu Glu Phe Gly Glu Gly Leu Trp
        195                 200                 205

ATT CTC GAC CAG AAG ATA ATG AGT TGG TTC TCG GAG CAG TAC TTC TCC        672
Ile Leu Asp Gln Lys Ile Met Ser Trp Phe Ser Glu Gln Tyr Phe Ser
    210                 215                 220

AGA GAG GAA GAT AAG TTC AAG CCC CTC GCC TCC GTA ATC TTT GCG GAC        720
Arg Glu Glu Asp Lys Phe Asn Pro Leu Ala Ser Val Ile Phe Ala Asp
225                 230                 235                 240

CTT GAG AAC CTA CCT CCT GCG CTG ATC ATA ACC GCC GAA TAC GAC CCG        768
Leu Glu Asn Leu Pro Pro Ala Leu Ile Ile Thr Ala Glu Tyr Asp Pro
                245                 250                 255

CTG AGA GAT GAA GGA GAA GTT TTC GGG CAG ATG CTG AGA AGA GCC GGT        816
Leu Arg Asp Glu Gly Glu Val Phe Gly Gln Met Leu Arg Arg Ala Gly
            260                 265                 270

GTT GAG GCG AGC ATC GTC AGA TAC AGA GGC GTG CTT CAC GGA TTC ATC        864
Val Glu Ala Ser Ile Val Arg Tyr Arg Gly Val Leu His Gly Phe Ile
        275                 280                 285

AAT TAC TAT CCC GTG CTG AAG GCT GCG AGG GAT GCG ATA AAC CAG ATT        912
Asn Tyr Tyr Pro Val Leu Lys Ala Ala Arg Asp Ala Ile Asn Gln Ile
    290                 295                 300

GCC GCT CTT CTT GTG TTC GAC TAG                                        936
Ala Ala Leu Leu Val Phe Asp
305                 310
```

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 918 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:

(A) NAME/KEY: Coding Sequence
(B) LOCATION: 1...915

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

```
ATG CCC CTA GAT CCT AGA ATT AAA AAG TTA CTA GAA TCA GCT CTT ACT        48
Met Pro Leu Asp Pro Arg Ile Lys Lys Leu Leu Glu Ser Ala Leu Thr
 1               5                   10                  15

ATA CCA ATT GGT AAA GCC CCA GTA GAA GAG GTA AGA AAG ATA TTT AGG        96
Ile Pro Ile Gly Lys Ala Pro Val Glu Glu Val Arg Lys Ile Phe Arg
             20                  25                  30

CAA TTA GCG TCG GCA GCT CCC AAA GTC GAA GTT GGA AAA GTA GAA GAT       144
Gln Leu Ala Ser Ala Ala Pro Lys Val Glu Val Gly Lys Val Glu Asp
         35                  40                  45

ATA AAA ATA CCA GGC AGT GAA ACC GTT ATA AAC GCT AGA GTG TAT TTT       192
Ile Lys Ile Pro Gly Ser Glu Thr Val Ile Asn Ala Arg Val Tyr Phe
     50                  55                  60

CCG AAG AGT AGC GGT CCT TAT GGT GTT CTA GTG TAT CTT CAT GGA GGC       240
Pro Lys Ser Ser Gly Pro Tyr Gly Val Leu Val Tyr Leu His Gly Gly
65                   70                  75                  80
GGT TTT GTA ATA GGC GAT GTG GAA TCT TAT GAC CCA TTA TGT AGA GCA       288
Gly Phe Val Ile Gly Asp Val Glu Ser Tyr Asp Pro Leu Cys Arg Ala
                 85                  90                  95

ATT ACA AAT GCG TGC AAT TGC GTT GTA GTA TCA GTG GAC TAT AGG TTA       336
Ile Thr Asn Ala Cys Asn Cys Val Val Val Ser Val Asp Tyr Arg Leu
            100                 105                 110

GCT CCA GAA TAC AAG TTT CCT TCT GCA GTT ATC GAT TCA TTT GAC GCT       384
Ala Pro Glu Tyr Lys Phe Pro Ser Ala Val Ile Asp Ser Phe Asp Ala
        115                 120                 125

ACT AAT TGG GTT TAT AAC AAT TTA GAT AAA TTT GAT GGA AAG ATG GGA       432
Thr Asn Trp Val Tyr Asn Asn Leu Asp Lys Phe Asp Gly Lys Met Gly
    130                 135                 140

GTT GCG ATT GCG GGA GAT AGT GCT GGA GGA AAT TTG GCA GCG GTT GTA       480
Val Ala Ile Ala Gly Asp Ser Ala Gly Gly Asn Leu Ala Ala Val Val
145                 150                 155                 160

GCT CTT CTT TCA AAG GGT AAA ATT AAT TTG AAG TAT CAA ATA CTG GTT       528
Ala Leu Leu Ser Lys Gly Lys Ile Asn Leu Lys Tyr Gln Ile Leu Val
                165                 170                 175
```

```
TAC CCA GCG GTA AGT TTA GAT AAC GTT TCA AGA TCC ATG ATA GAG TAC         576
Tyr Pro Ala Val Ser Leu Asp Asn Val Ser Arg Ser Met Ile Glu Tyr
            180                 185                 190

TCT GAT GGG TTC TTC CTT ACC AGA GAG CAT ATA GAG TGG TTC GGT TCT         624
Ser Asp Gly Phe Phe Leu Thr Arg Glu His Ile Glu Trp Phe Gly Ser
        195                 200                 205

CAA TAC TTA CGA AGC CCT GCA GAT TTG CTA GAC TTT AGG TTC TCT CCA         672
Gln Tyr Leu Arg Ser Pro Ala Asp Leu Leu Asp Phe Arg Phe Ser Pro
    210                 215                 220

ATT CTG GCG CAA GAT TTC AAC GGA TTA CCT CCA GCC TTG ATA ATA ACA         720
Ile Leu Ala Gln Asp Phe Asn Gly Leu Pro Pro Ala Leu Ile Ile Thr
225                 230                 235                 240

GCA GAA TAC GAT CCA CTA AGG GAT CAA GGA GAA GCG TAT GCA AAT AAA         768
Ala Glu Tyr Asp Pro Leu Arg Asp Gln Gly Glu Ala Tyr Ala Asn Lys
                245                 250                 255

CTA CTA CAA GCT GGA GTC TCA GTT ACT AGT GTG AGA TTT AAC AAC GTT         816
Leu Leu Gln Ala Gly Val Ser Val Thr Ser Val Arg Phe Asn Asn Val
            260                 265                 270

ATA CAC GGA TTC CTC TCA TTC TTT CCG TTG ATG GAG CAA GGA AGA GAT         864
Ile His Gly Phe Leu Ser Phe Phe Pro Leu Met Glu Gln Gly Arg Asp
        275                 280                 285

GCT ATA GGT CTG ATA GGG TCT GTG TTA AGA CGA GTA TTT TAT GAT AAA         912
Ala Ile Gly Leu Ile Gly Ser Val Leu Arg Arg Val Phe Tyr Asp Lys
    290                 295                 300

ATT TAA                                                                 918
Ile
305
```

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 184 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

```
Met Ser Leu Asn Lys His Ser Trp Met Asp Met Ile Ile Phe Ile Leu
 1               5                  10                  15
Ser Phe Ser Phe Pro Leu Thr Met Ile Ala Leu Ala Ile Ser Met Ser
            20                  25                  30
Ser Trp Phe Asn Ile Trp Asn Asn Ala Leu Ser Asp Leu Gly His Ala
        35                  40                  45
```

```
Val Lys Ser Ser Val Ala Pro Ile Phe Asn Leu Gly Leu Ala Ile Gly
     50                  55                  60
Gly Ile Leu Ile Val Ile Val Gly Leu Arg Asn Leu Tyr Ser Trp Ser
 65                  70                  75                  80
Arg Val Lys Gly Ser Leu Ile Ile Ser Met Gly Val Phe Leu Asn Leu
                 85                  90                  95
Ile Gly Val Phe Asp Glu Val Tyr Gly Trp Ile His Phe Leu Val Ser
            100                 105                 110
Val Leu Phe Phe Leu Ser Ile Ile Ala Tyr Phe Ile Ala Ile Ser Ile
        115                 120                 125
Leu Asp Lys Ser Trp Ile Ala Val Leu Leu Ile Ile Gly His Ile Ala
    130                 135                 140
Met Trp Tyr Leu His Phe Ala Ser Glu Ile Pro Arg Gly Ala Ala Ile
145                 150                 155                 160
Pro Glu Leu Leu Ala Val Phe Ser Phe Leu Pro Phe Tyr Ile Arg Asp
                165                 170                 175
Tyr Phe Lys Ser Tyr Thr Lys Arg
            180
```

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 346 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

```
Met Lys Leu Leu Glu Pro Thr Asn Thr Ser Tyr Thr Leu Leu Gln Asp
  1               5                  10                  15
Leu Ala Leu His Phe Ala Phe Tyr Trp Phe Leu Ala Val Tyr Thr Trp
             20                  25                  30
Leu Pro Gly Val Leu Val Arg Gly Val Ala Val Asp Thr Gly Val Ala
         35                  40                  45
Arg Val Pro Gly Leu Gly Arg Arg Gly Lys Arg Leu Leu Leu Ala Ala
     50                  55                  60
Val Ala Val Leu Ala Leu Val Val Ser Val Val Val Pro Ala Tyr Val
 65                  70                  75                  80
Ala Tyr Ser Ser Leu His Pro Glu Ser Cys Arg Pro Val Ala Pro Glu
                 85                  90                  95
Gly Leu Thr Tyr Lys Glu Phe Ser Val Thr Ala Glu Asp Gly Leu Val
            100                 105                 110
Val Arg Gly Trp Val Leu Gly Pro Gly Ala Gly Gly Asn Pro Val Phe
        115                 120                 125
Val Leu Met His Gly Tyr Thr Gly Cys Arg Ser Ala Pro Tyr Met Ala
    130                 135                 140
Val Leu Ala Arg Glu Leu Val Glu Trp Gly Tyr Pro Val Val Val Phe
145                 150                 155                 160
Asp Phe Arg Gly His Gly Glu Ser Gly Gly Ser Thr Thr Ile Gly Pro
                165                 170                 175
```

```
Arg Glu Val Leu Asp Ala Arg Ala Val Val Gly Tyr Val Ser Glu Arg
            180                 185                 190
Phe Pro Gly Arg Arg Ile Ile Leu Val Gly Phe Ser Met Gly Gly Ala
        195                 200                 205
Val Ala Ile Val Glu Gly Ala Gly Asp Pro Arg Val Tyr Ala Val Ala
    210                 215                 220
Ala Asp Ser Pro Tyr Tyr Arg Leu Arg Asp Val Ile Pro Arg Trp Leu
225                 230                 235                 240
Glu Tyr Lys Thr Pro Leu Pro Gly Trp Val Gly Val Leu Ala Gly Phe
                245                 250                 255
Tyr Gly Arg Leu Met Ala Gly Val Asp Leu Gly Phe Gly Pro Ala Gly
            260                 265                 270
Val Glu Arg Val Asp Lys Pro Leu Leu Val Val Tyr Gly Pro Arg Asp
        275                 280                 285
Pro Leu Val Thr Arg Asp Glu Ala Arg Ser Leu Ala Ser Arg Ser Pro
    290                 295                 300
Cys Gly Arg Leu Val Glu Val Pro Gly Ala Gly His Val Glu Ala Val
305                 310                 315                 320
Asp Val Leu Gly Pro Gly Arg Tyr Ala Asp Met Leu Ile Glu Leu Ala
                325                 330                 335
His Glu Glu Cys Pro Pro Gly Ala Gly Gly
            340                 345
```

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 262 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

```
Met Pro Tyr Val Arg Asn Gly Gly Val Asn Ile Tyr Tyr Glu Leu Val
 1               5                  10                  15
Asp Gly Pro Glu Pro Pro Ile Val Phe Val His Gly Trp Thr Ala Asn
             20                  25                  30
Met Asn Phe Trp Lys Glu Gln Arg Arg Tyr Phe Ala Gly Arg Asn Met
         35                  40                  45
Met Leu Phe Val Asp Asn Arg Gly His Gly Arg Ser Asp Lys Pro Leu
     50                  55                  60
Gly Tyr Asp Phe Tyr Arg Phe Glu Asn Phe Ile Ser Asp Leu Asp Ala
 65                  70                  75                  80
Val Val Arg Glu Thr Gly Val Glu Lys Phe Val Leu Val Gly His Ser
                 85                  90                  95
Phe Gly Thr Met Ile Ser Met Lys Tyr Cys Ser Glu Tyr Arg Asn Arg
            100                 105                 110
Val Leu Ala Leu Ile Leu Ile Gly Gly Gly Ser Arg Ile Lys Leu Leu
        115                 120                 125
His Arg Ile Gly Tyr Pro Leu Ala Lys Ile Leu Ala Ser Ile Ala Tyr
    130                 135                 140
```

```
Lys Lys Ser Ser Arg Leu Val Ala Asp Leu Ser Phe Gly Lys Asn Ala
145                 150                 155                 160
Gly Glu Leu Lys Glu Trp Gly Trp Lys Gln Ala Met Asp Tyr Thr Pro
                165                 170                 175
Ser Tyr Val Ala Met Tyr Thr Tyr Arg Thr Leu Thr Lys Val Asn Leu
            180                 185                 190
Glu Asn Ile Leu Glu Lys Ile Asp Cys Pro Thr Leu Ile Ile Val Gly
        195                 200                 205
Glu Glu Asp Ala Leu Leu Pro Val Ser Lys Ser Val Glu Leu Ser Arg
    210                 215                 220
Arg Ile Glu Asn Ser Lys Leu Val Ile Ile Pro Asn Ser Gly His Cys
225                 230                 235                 240
Val Met Leu Glu Ser Pro Ser Glu Val Asn Arg Ala Met Asp Glu Phe
                245                 250                 255
Ile Ser Ser Ala Gln Phe
            260
```

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 251 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

```
Leu Arg Leu Arg Lys Phe Glu Glu Ile Asn Leu Val Leu Ser Gly Gly
 1               5                  10                  15
Ala Ala Lys Gly Ile Ala His Ile Gly Val Leu Lys Ala Ile Asn Glu
            20                  25                  30
Leu Gly Ile Arg Val Arg Ala Leu Ser Gly Val Ser Ala Gly Ala Ile
        35                  40                  45
Val Ser Val Phe Tyr Ala Ser Gly Tyr Ser Pro Glu Gly Met Phe Ser
    50                  55                  60
Leu Leu Lys Arg Val Asn Trp Leu Lys Leu Phe Lys Phe Lys Pro Pro
65                  70                  75                  80
Leu Lys Gly Leu Ile Gly Trp Glu Lys Ala Ile Arg Phe Leu Glu Glu
                85                  90                  95
Val Leu Pro Tyr Arg Arg Ile Glu Lys Leu Glu Ile Pro Thr Tyr Ile
            100                 105                 110
Cys Ala Thr Asp Leu Tyr Ser Gly Arg Ala Leu Tyr Leu Ser Glu Gly
        115                 120                 125
Ser Leu Ile Pro Ala Leu Leu Gly Ser Cys Ala Ile Pro Gly Ile Phe
    130                 135                 140
Glu Pro Val Glu Tyr Lys Asn Tyr Leu Leu Val Asp Gly Gly Ile Val
145                 150                 155                 160
Asn Asn Leu Pro Val Glu Pro Phe Gln Glu Ser Gly Ile Pro Thr Val
                165                 170                 175
Cys Val Asp Val Leu Pro Ile Glu Pro Glu Lys Asp Ile Lys Asn Ile
            180                 185                 190
Leu His Ile Leu Leu Arg Ser Phe Phe Leu Ala Val Arg Ser Asn Ser
        195                 200                 205
```

```
Glu Lys Arg Lys Glu Phe Cys Asp Leu Val Ile Val Pro Glu Leu Glu
    210                 215                 220
Glu Phe Thr Pro Leu Asp Val Arg Lys Ala Asp Gln Ile Met Glu Arg
225                 230                 235                 240
Gly Tyr Ile Lys Ala Leu Glu Val Leu Ser Glu
                245                 250
```

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 297 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

```
Met Phe Asn Ile Asn Val Phe Val Asn Ile Ser Trp Leu Tyr Phe Ser
1               5                   10                  15
Gly Ile Val Met Lys Thr Val Glu Glu Tyr Ala Leu Leu Glu Thr Gly
            20                  25                  30
Val Arg Val Phe Tyr Arg Cys Val Ile Pro Glu Lys Ala Phe Asn Thr
        35                  40                  45
Leu Ile Ile Gly Ser His Gly Leu Gly Ala His Ser Gly Ile Tyr Ile
    50                  55                  60
Ser Val Ala Glu Glu Phe Ala Arg His Gly Phe Gly Phe Cys Met His
65                  70                  75                  80
Asp Gln Arg Gly His Gly Arg Thr Ala Ser Asp Arg Glu Arg Gly Tyr
                85                  90                  95
Val Glu Gly Phe His Asn Phe Ile Glu Asp Met Lys Ala Phe Ser Asp
            100                 105                 110
Tyr Ala Lys Trp Arg Val Gly Gly Asp Glu Ile Ile Leu Leu Gly His
        115                 120                 125
Ser Met Gly Gly Leu Ile Ala Leu Leu Thr Val Ala Thr Tyr Lys Glu
    130                 135                 140
Ile Ala Lys Gly Val Ile Ala Leu Ala Pro Ala Leu Gln Ile Pro Leu
145                 150                 155                 160
Thr Pro Ala Arg Arg Leu Val Leu Ser Leu Ala Ser Arg Leu Ala Pro
                165                 170                 175
His Ser Lys Ile Thr Leu Gln Arg Arg Leu Pro Gln Lys Pro Glu Gly
            180                 185                 190
Phe Gln Arg Ala Lys Asp Ile Glu Tyr Ser Leu Ser Glu Ile Ser Val
        195                 200                 205
Lys Leu Val Asp Glu Met Ile Lys Ala Ser Ser Met Phe Trp Thr Ile
    210                 215                 220
Ala Gly Glu Ile Asn Thr Pro Val Leu Leu Ile His Gly Glu Lys Asp
225                 230                 235                 240
Asn Val Ile Pro Pro Glu Ala Ser Lys Lys Ala Tyr Gln Leu Ile Pro
                245                 250                 255
Ser Phe Pro Lys Glu Leu Lys Ile Tyr Pro Asp Leu Gly His Asn Leu
            260                 265                 270
```

```
Phe Phe Glu Pro Gly Ala Val Lys Ile Val Thr Asp Ile Val Glu Trp
        275                 280                 285
Val Lys Asn Leu Pro Arg Glu Asn Pro
    290                 295
```

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 262 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

```
Met Glu Val Tyr Lys Ala Lys Phe Gly Glu Ala Lys Leu Gly Trp Val
 1               5                  10                  15
Val Leu Val His Gly Leu Gly Glu His Ser Gly Arg Tyr Gly Arg Leu
             20                  25                  30
Ile Lys Glu Leu Asn Tyr Ala Gly Phe Gly Val Tyr Thr Phe Asp Trp
         35                  40                  45
Pro Gly His Gly Lys Ser Pro Gly Lys Arg Gly His Thr Ser Val Glu
     50                  55                  60
Glu Ala Met Glu Ile Ile Asp Ser Ile Ile Glu Glu Ile Arg Glu Lys
 65                  70                  75                  80
Pro Phe Leu Phe Gly His Ser Leu Gly Gly Leu Thr Val Ile Arg Tyr
                 85                  90                  95
Ala Glu Thr Arg Pro Asp Lys Ile Arg Gly Leu Ile Ala Ser Ser Pro
            100                 105                 110
Ala Leu Ala Lys Ser Pro Glu Thr Pro Gly Phe Met Val Ala Leu Ala
        115                 120                 125
Lys Phe Leu Gly Lys Ile Ala Pro Gly Val Val Leu Ser Asn Gly Ile
    130                 135                 140
Lys Pro Glu Leu Leu Ser Arg Asn Arg Asp Ala Val Arg Arg Tyr Val
145                 150                 155                 160
Glu Asp Pro Leu Val His Asp Arg Ile Ser Ala Lys Leu Gly Arg Ser
                165                 170                 175
Ile Phe Val Asn Met Glu Leu Ala His Arg Glu Ala Asp Lys Ile Lys
            180                 185                 190
Val Pro Ile Leu Leu Leu Ile Gly Thr Gly Asp Val Ile Thr Pro Pro
        195                 200                 205
Glu Gly Ser Arg Arg Leu Phe Glu Glu Leu Ala Val Glu Asn Lys Thr
    210                 215                 220
Leu Arg Glu Phe Glu Gly Ala Tyr His Glu Ile Phe Glu Asp Pro Glu
225                 230                 235                 240
Trp Ala Glu Glu Phe His Glu Thr Ile Val Lys Trp Leu Val Glu Lys
                245                 250                 255
Ser Tyr Ser Ser Ala Gln
            260
```

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 249 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

```
Leu Ile Gly Asn Leu Lys Leu Lys Arg Phe Glu Glu Val Asn Leu Val
 1               5                  10                  15
Leu Ser Gly Gly Ala Ala Lys Gly Ile Ala His Ile Gly Val Leu Lys
             20                  25                  30
Ala Leu Glu Glu Leu Gly Ile Lys Val Lys Arg Leu Ser Gly Val Ser
         35                  40                  45
Ala Gly Ala Ile Val Ser Val Phe Tyr Ala Ser Gly Tyr Thr Pro Asp
     50                  55                  60
Glu Met Leu Lys Leu Leu Lys Glu Val Asn Trp Leu Lys Leu Phe Lys
 65                  70                  75                  80
Phe Lys Thr Pro Lys Met Gly Leu Met Gly Trp Glu Lys Ala Ala Glu
                     85                  90                  95
Phe Leu Glu Lys Glu Leu Gly Val Lys Arg Leu Glu Asp Leu Asn Ile
            100                 105                 110
Pro Thr Tyr Leu Cys Ser Ala Asp Leu Tyr Thr Gly Lys Ala Leu Tyr
        115                 120                 125
Phe Gly Arg Gly Asp Leu Ile Pro Val Leu Leu Gly Ser Cys Ser Ile
    130                 135                 140
Pro Gly Ile Phe Glu Pro Val Glu Tyr Glu Asn Phe Leu Leu Val Asp
145                 150                 155                 160
Gly Gly Ile Val Asn Asn Leu Pro Val Glu Pro Leu Glu Lys Phe Lys
                165                 170                 175
Glu Pro Ile Ile Gly Val Asp Val Leu Pro Ile Thr Gln Glu Arg Lys
            180                 185                 190
Ile Lys Asn Ile Leu His Ile Leu Ile Arg Ser Phe Phe Leu Ala Val
        195                 200                 205
Arg Ser Asn Ser Glu Lys Arg Lys Glu Phe Cys Asn Val Val Ile Glu
    210                 215                 220
Pro Pro Leu Glu Glu Phe Ser Pro Leu Asp Val Asn Lys Ala Asp Glu
225                 230                 235                 240
Ile Phe Cys Gly Asp Met Arg Ala Leu
                    245
```

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 338 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

```
Met Pro Ala Asn Asp Ser Pro Thr Ile Asp Phe Asn Pro Arg Gly Ile
1               5                   10                  15
Leu Arg Asn Ala His Ala Gln Val Ile Leu Ala Thr Ser Gly Leu Arg
            20                  25                  30
Lys Ala Phe Leu Lys Arg Thr His Lys Ser Tyr Leu Ser Thr Ala Gln
        35                  40                  45
Trp Leu Glu Leu Asp Ala Gly Asn Gly Val Thr Leu Ala Gly Glu Leu
    50                  55                  60
Asn Thr Ala Pro Ala Thr Ala Ser Ser Ser His Pro Ala His Lys Asn
65                  70                  75                  80
Thr Leu Val Ile Val Leu His Gly Trp Glu Gly Ser Ser Gln Ser Ala
                85                  90                  95
Tyr Ala Thr Ser Ala Gly Ser Thr Leu Phe Asp Asn Gly Phe Asp Thr
            100                 105                 110
Phe Arg Leu Asn Phe Arg Asp His Gly Asp Thr Tyr His Leu Asn Arg
        115                 120                 125
Gly Ile Phe Asn Ser Ser Leu Ile Asp Glu Val Val Gly Ala Val Lys
    130                 135                 140
Ala Ile Gln Gln Gln Thr Asp Tyr Asp Lys Tyr Cys Leu Met Gly Phe
145                 150                 155                 160
Ser Leu Gly Gly Asn Phe Ala Leu Arg Val Ala Val Arg Glu Gln His
                165                 170                 175
Leu Ala Lys Pro Leu Ala Gly Val Leu Ala Val Cys Pro Val Leu Asp
            180                 185                 190
Pro Ala His Thr Met Met Ala Leu Asn Arg Gly Ala Phe Phe Tyr Gly
        195                 200                 205
Arg Tyr Phe Ala His Lys Trp Lys Arg Ser Leu Thr Ala Lys Leu Ala
    210                 215                 220
Ala Phe Pro Asp Tyr Lys Tyr Gly Lys Asp Leu Lys Ser Ile His Thr
225                 230                 235                 240
Leu Asp Glu Leu Asn Asn Tyr Phe Ile Pro Arg Tyr Thr Gly Phe Asn
                245                 250                 255
Ser Val Ser Glu Tyr Phe Lys Ser Tyr Thr Leu Thr Gly Gln Lys Leu
            260                 265                 270
Ala Phe Leu Asn Cys Pro Ser Tyr Ile Leu Ala Ala Gly Asp Asp Pro
        275                 280                 285
Ile Ile Pro Ala Ser Asp Phe Gln Lys Ile Ala Lys Pro Ala Asn Leu
    290                 295                 300
His Ile Thr Val Thr Gln Gln Gly Ser His Cys Ala Tyr Leu Glu Asn
305                 310                 315                 320
Leu His Lys Pro Ser Ala Ala Asp Lys Tyr Ala Val Lys Leu Phe Gly
                325                 330                 335
Ala Cys
```

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 311 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

```
Met Leu Asp Met Pro Ile Asp Pro Val Tyr Tyr Gln Leu Ala Glu Tyr
1               5                   10                  15
Phe Asp Ser Leu Pro Lys Phe Asp Gln Phe Ser Ser Ala Arg Glu Tyr
            20                  25                  30
Arg Glu Ala Ile Asn Arg Ile Tyr Glu Glu Arg Asn Arg Gln Leu Ser
        35                  40                  45
Gln His Glu Arg Val Glu Arg Val Glu Asp Arg Thr Ile Lys Gly Arg
    50                  55                  60
Asn Gly Asp Ile Arg Val Arg Val Tyr Gln Gln Lys Pro Asp Ser Pro
65                  70                  75                  80
Val Leu Val Tyr Tyr His Gly Gly Gly Phe Val Ile Cys Ser Ile Glu
                85                  90                  95
Ser His Asp Ala Leu Cys Arg Arg Ile Ala Arg Leu Ser Asn Ser Thr
            100                 105                 110
Val Val Ser Val Asp Tyr Arg Leu Ala Pro Glu His Lys Phe Pro Ala
        115                 120                 125
Ala Val Tyr Asp Cys Tyr Asp Ala Thr Lys Trp Val Ala Glu Asn Ala
    130                 135                 140
Glu Glu Leu Arg Ile Asp Pro Ser Lys Ile Phe Val Gly Gly Asp Ser
145                 150                 155                 160
Ala Gly Gly Asn Leu Ala Ala Ala Val Ser Ile Met Ala Arg Asp Ser
                165                 170                 175
Gly Glu Asp Phe Ile Lys His Gln Ile Leu Ile Tyr Pro Val Val Asn
            180                 185                 190
Phe Val Ala Pro Thr Pro Ser Leu Leu Glu Phe Gly Glu Gly Leu Trp
        195                 200                 205
Ile Leu Asp Gln Lys Ile Met Ser Trp Phe Ser Glu Gln Tyr Phe Ser
    210                 215                 220
Arg Glu Glu Asp Lys Phe Asn Pro Leu Ala Ser Val Ile Phe Ala Asp
225                 230                 235                 240
Leu Glu Asn Leu Pro Pro Ala Leu Ile Ile Thr Ala Glu Tyr Asp Pro
                245                 250                 255
Leu Arg Asp Glu Gly Glu Val Phe Gly Gln Met Leu Arg Arg Ala Gly
            260                 265                 270
Val Glu Ala Ser Ile Val Arg Tyr Arg Gly Val Leu His Gly Phe Ile
        275                 280                 285
Asn Tyr Tyr Pro Val Leu Lys Ala Ala Arg Asp Ala Ile Asn Gln Ile
    290                 295                 300
Ala Ala Leu Leu Val Phe Asp
305                 310
```

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 305 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

```
Met Pro Leu Asp Pro Arg Ile Lys Lys Leu Leu Glu Ser Ala Leu Thr
 1               5                  10                  15
Ile Pro Ile Gly Lys Ala Pro Val Glu Glu Val Arg Lys Ile Phe Arg
            20                  25                  30
Gln Leu Ala Ser Ala Ala Pro Lys Val Glu Val Gly Lys Val Glu Asp
        35                  40                  45
Ile Lys Ile Pro Gly Ser Glu Thr Val Ile Asn Ala Arg Val Tyr Phe
    50                  55                  60
Pro Lys Ser Ser Gly Pro Tyr Gly Val Leu Val Tyr Leu His Gly Gly
65                  70                  75                  80
Gly Phe Val Ile Gly Asp Val Glu Ser Tyr Asp Pro Leu Cys Arg Ala
                85                  90                  95
Ile Thr Asn Ala Cys Asn Cys Val Val Val Ser Val Asp Tyr Arg Leu
            100                 105                 110
Ala Pro Glu Tyr Lys Phe Pro Ser Ala Val Ile Asp Ser Phe Asp Ala
        115                 120                 125
Thr Asn Trp Val Tyr Asn Asn Leu Asp Lys Phe Asp Gly Lys Met Gly
    130                 135                 140
Val Ala Ile Ala Gly Asp Ser Ala Gly Gly Asn Leu Ala Ala Val Val
145                 150                 155                 160
Ala Leu Leu Ser Lys Gly Lys Ile Asn Leu Lys Tyr Gln Ile Leu Val
                165                 170                 175
Tyr Pro Ala Val Ser Leu Asp Asn Val Ser Arg Ser Met Ile Glu Tyr
            180                 185                 190
Ser Asp Gly Phe Phe Leu Thr Arg Glu His Ile Glu Trp Phe Gly Ser
        195                 200                 205
Gln Tyr Leu Arg Ser Pro Ala Asp Leu Leu Asp Phe Arg Phe Ser Pro
    210                 215                 220
Ile Leu Ala Gln Asp Phe Asn Gly Leu Pro Pro Ala Leu Ile Ile Thr
225                 230                 235                 240
Ala Glu Tyr Asp Pro Leu Arg Asp Gln Gly Glu Ala Tyr Ala Asn Lys
                245                 250                 255
Leu Leu Gln Ala Gly Val Ser Val Thr Ser Val Arg Phe Asn Asn Val
            260                 265                 270
Ile His Gly Phe Leu Ser Phe Phe Pro Leu Met Glu Gln Gly Arg Asp
        275                 280                 285
Ala Ile Gly Leu Ile Gly Ser Val Leu Arg Arg Val Phe Tyr Asp Lys
    290                 295                 300
Ile
305
```

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 605 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:

(A) NAME/KEY: Coding Sequence
(B) LOCATION: 1...603

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

```
ATG AAG GTT AAA CAC GTT ATT GTT TTA CAT GGC TTA TAT ATG TCT GGC       48
Met Lys Val Lys His Val Ile Val Leu His Gly Leu Tyr Met Ser Gly
 1               5                  10                  15

TTG GTG ATG CGC CCG TTA TGT TCG CGT CTA GAA GAG TCG GGG GTT AAA       96
Leu Val Met Arg Pro Leu Cys Ser Arg Leu Glu Glu Ser Gly Val Lys
            20                  25                  30

GTT TTA AAC TTA ACC TAC AAT ACT CGA GAC CCT AAT CGA GAT GCT ATT      144
Val Leu Asn Leu Thr Tyr Asn Thr Arg Asp Pro Asn Arg Asp Ala Ile
        35                  40                  45

TTT ACG CAA ATA GAT GAG TTT ATT AGC AAT GAG CCT TCT GCT TTA GTG      192
Phe Thr Gln Ile Asp Glu Phe Ile Ser Asn Glu Pro Ser Ala Leu Val
    50                  55                  60

TGT CAC TCT ATG GGG GGC TTA GTT GCT CGC GCC TAT TTA GAG GCA AAC      240
Cys His Ser Met Gly Gly Leu Val Ala Arg Ala Tyr Leu Glu Ala Asn
65                  70                  75                      80

TCA GCG CCA AGT CAT CAT GTT GAA AAG GTA ATC ACC TTA GGA ACG CCA      288
Ser Ala Pro Ser His His Val Glu Lys Val Ile Thr Leu Gly Thr Pro
                85                  90                  95

CAT ACT GGC AGC CAT ATT GCT GAA AAA ATG CAG CAA AAA GGG TTC GAG      336
His Thr Gly Ser His Ile Ala Glu Lys Met Gln Gln Lys Gly Phe Glu
            100                 105                 110

CTA TTA TTA AAA AAT AGC GTT GAG TTT TTA CTC TCT AAG AAT GGT GAT      384
Leu Leu Leu Lys Asn Ser Val Glu Phe Leu Leu Ser Lys Asn Gly Asp
        115                 120                 125

TGG CCT TTT AAA GCC AAG CTA TAT AGC ATT GCC GGC GAC TTA CCG ATT      432
Trp Pro Phe Lys Ala Lys Leu Tyr Ser Ile Ala Gly Asp Leu Pro Ile
    130                 135                 140

GGC TTA ATG CCA CTC ATT GTA AAA GGC AGC CGC TCT GAT GGC ACT GTA      480
Gly Leu Met Pro Leu Ile Val Lys Gly Ser Arg Ser Asp Gly Thr Val
145                 150                 155                     160

TTG CTA GAT GAA ACC AAG CTA AAG GGT ATG GCT GAA CAC AAG GTG TTT      528
Leu Leu Asp Glu Thr Lys Leu Lys Gly Met Ala Glu His Lys Val Phe
                165                 170                 175

CAT TTA AGC CAT ACA AGT ATG ATT TAC TCT CGC CAA GTC GTT AAT TAT      576


His Leu Ser His Thr Ser Met Ile Tyr Ser Arg Gln Val Val Asn Tyr
            180                 185                 190

ATT CTT GAG CGC TTG AAC GAG GAC ATT TA                               605
Ile Leu Glu Arg Leu Asn Glu Asp Ile
        195                 200
```

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 779 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:

(A) NAME/KEY: Coding Sequence
(B) LOCATION: 1...777

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

```
ATG ATA AAA AAC TTC GAC AGA GAA AAT TCT AGC TTA GTA CTG TCC GGT       48
Met Ile Lys Asn Phe Asp Arg Glu Asn Ser Ser Leu Val Leu Ser Gly
  1               5                  10                  15

GGT GGT GCT CTG GGT ATT GCT CAC TTG GGT GTA CTG CAT GAC CTT GAA       96
Gly Gly Ala Leu Gly Ile Ala His Leu Gly Val Leu His Asp Leu Glu
             20                  25                  30

AAA CAA AAT ATT GTA CCA AAT GAA ATT GTT GGT ACA AGT ATG GGT GGT      144
Lys Gln Asn Ile Val Pro Asn Glu Ile Val Gly Thr Ser Met Gly Gly
         35                  40                  45

ATC ATT GGT GCA TCT ATG GCT ATC GGG ATG AAA GAG AAA GAA ATA CTC      192
Ile Ile Gly Ala Ser Met Ala Ile Gly Met Lys Glu Lys Glu Ile Leu
     50                  55                  60

GAA GAA ATC AAA AAC TTT TCC AAT GTC TTC AAC TGG ATA AAA TTC TCT      240
Glu Glu Ile Lys Asn Phe Ser Asn Val Phe Asn Trp Ile Lys Phe Ser
 65                  70                  75                  80

TTT TCC GGT AAT TCT GTT GTC GAT AAC GAG AAG ATC GCT AAG ATA TTT      288
Phe Ser Gly Asn Ser Val Val Asp Asn Glu Lys Ile Ala Lys Ile Phe
                 85                  90                  95

GAT ACT CTT TTT AAA GAC AGA AAG ATG ACA GAT ACG GTG ATC CCT CTT      336
Asp Thr Leu Phe Lys Asp Arg Lys Met Thr Asp Thr Val Ile Pro Leu
            100                 105                 110

AAA CTC ATC GCT ACA AAC TTA CAT AAT GGA CAT AAA AAA GTA TTT ACT      384
Lys Leu Ile Ala Thr Asn Leu His Asn Gly His Lys Lys Val Phe Thr
        115                 120                 125
```

```
GCT TCG GAT GAT GTA CTG ATC AAA GAT GCA ATA CTC TCA ACA ATG GCA       432
Ala Ser Asp Asp Val Leu Ile Lys Asp Ala Ile Leu Ser Thr Met Ala
    130                 135                 140

ATA CCC GGT GTA TTT GAA GAA CAT ATT ATT GAT GGT GAA ACC TAT GGC       480
Ile Pro Gly Val Phe Glu Glu His Ile Ile Asp Gly Glu Thr Tyr Gly
145                 150                 155                 160

GAC GGT TTT CTT TGT GAA AAC CTT GGT GTG AAT GAG GCA ACA TTC AAT       528
Asp Gly Phe Leu Cys Glu Asn Leu Gly Val Asn Glu Ala Thr Phe Asn
                165                 170                 175

GAT GTT TTA GCT GTA GAT GTC ATG GGT GAG AAC TCT TTT GAA AAA GCA       576
Asp Val Leu Ala Val Asp Val Met Gly Glu Asn Ser Phe Glu Lys Ala
            180                 185                 190

ATG CCG GAC AAC TTC TTT AAA ACA TCA AAT GTT TTA GAA ATG TTT GAA       624
Met Pro Asp Asn Phe Phe Lys Thr Ser Asn Val Leu Glu Met Phe Glu
        195                 200                 205

AAA TCA ATG CGA CTT TTT ATT TAC AAC CAG ACA CAG ACA CAT ATT AAA       672
Lys Ser Met Arg Leu Phe Ile Tyr Asn Gln Thr Gln Thr His Ile Lys
    210                 215                 220

AAT GCA AAT AAA AAT ATT TAT CTT ATT GAA CCC GTT ACC AAA GAG TAT       720
Asn Ala Asn Lys Asn Ile Tyr Leu Ile Glu Pro Val Thr Lys Glu Tyr
225                 230                 235                 240

AAA ACA TTT CAA TTT CAT AAA CAT AAA GAG ATA CGT GCT TTA GGC TTG       768
Lys Thr Phe Gln Phe His Lys His Lys Glu Ile Arg Ala Leu Gly Leu
                245                 250                 255

GGT TTA CTG TG                                                        779
Gly Leu Leu
```

(2) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 905 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:

(A) NAME/KEY: Coding Sequence
(B) LOCATION: 1...903

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

```
ATG CCC CTA CAT CCA AAG GTA AAG AAA TTA CTT TCC CAG CTA CCT CCC       48
```

```
Met Pro Leu His Pro Lys Val Lys Lys Leu Leu Ser Gln Leu Pro Pro
 1               5                  10                  15

CAG GAC TTC TCC AGA AAC GTG CAG GAC CTG AGG AAG GCC TGG GAT TTA      96
Gln Asp Phe Ser Arg Asn Val Gln Asp Leu Arg Lys Ala Trp Asp Leu
            20                  25                  30

CCC TTC TCA GGG AGG AGG GAG ACC CTG AAG AGG GTT GAG GAC CTT GAG     144
Pro Phe Ser Gly Arg Arg Glu Thr Leu Lys Arg Val Glu Asp Leu Glu
        35                  40                  45

ATA CCC ACT AGG GAC GCA CGA ATC AGG GCC AGG GTC TAC ACC CCC TCA     192
Ile Pro Thr Arg Asp Ala Arg Ile Arg Ala Arg Val Tyr Thr Pro Ser
    50                  55                  60

AGT AAG GAA AAC TTA CCC GTC CTT GTT TAC TAT CAC GGC GGT GGC TTC     240
Ser Lys Glu Asn Leu Pro Val Leu Val Tyr Tyr His Gly Gly Gly Phe
65                  70                  75                  80

GTG TTC GGT AGC GTT GAC AGC TAC GAC GGC CTC GCA TCC CTT ATT GCC     288
Val Phe Gly Ser Val Asp Ser Tyr Asp Gly Leu Ala Ser Leu Ile Ala
                85                  90                  95

AAG GAA TCT GGG ATT GCG GTT ATC TCC GTG GAG TAT AGG CTC GCC CCT     336
Lys Glu Ser Gly Ile Ala Val Ile Ser Val Glu Tyr Arg Leu Ala Pro
            100                 105                 110

GAG CAC AAG TTC CCC ACC GCA GTC AAC GAC TCG TGG GAT GCG CTT CTC     384
Glu His Lys Phe Pro Thr Ala Val Asn Asp Ser Trp Asp Ala Leu Leu
        115                 120                 125

TGG ATC GCG GAG AAC GGA GGC AAG CTG GGG CTC GAC ACC TCG AGA CTT     432
Trp Ile Ala Glu Asn Gly Gly Lys Leu Gly Leu Asp Thr Ser Arg Leu
    130                 135                 140

GCC GTG GCT GGG GAT AGT GCT GGA GGA AAC CTG TCT GCC GTG GTG TCC     480
Ala Val Ala Gly Asp Ser Ala Gly Gly Asn Leu Ser Ala Val Val Ser
145                 150                 155                 160

CTC CTG GAC AGG GAC CAG GGT AAG GGA CTG GTT AGT TAT CAG GTC CTA     528
Leu Leu Asp Arg Asp Gln Gly Lys Gly Leu Val Ser Tyr Gln Val Leu
                165                 170                 175

ATC TAC CCA GCA GTG AAC ATG GTC GAT AAC TCC CCA TCC GTC AGG GAG     576
Ile Tyr Pro Ala Val Asn Met Val Asp Asn Ser Pro Ser Val Arg Glu
            180                 185                 190

TAC GGC GAG GGA TAC TTC CTC ACC AGG TCC ATG ATG AAC TGG TTC GGG     624
Tyr Gly Glu Gly Tyr Phe Leu Thr Arg Ser Met Met Asn Trp Phe Gly
        195                 200                 205

ACC ATG TAC TTC TCC TCT GGA AGG GAA GCG GTA TCC CCC TAC GCC TCT     672
Thr Met Tyr Phe Ser Ser Gly Arg Glu Ala Val Ser Pro Tyr Ala Ser
    210                 215                 220
```

```
CCA GCC TTG GCT GAC CTA CAT AAC CTC CCA CCC TCA CTG GTG ATC ACT       720
Pro Ala Leu Ala Asp Leu His Asn Leu Pro Pro Ser Leu Val Ile Thr
225                 230                 235                 240

GCA GAG TAT GAT CCC CTA AGG GAT CAG GGA GAG ACC TAC TCT CAC TCC       768
Ala Glu Tyr Asp Pro Leu Arg Asp Gln Gly Glu Thr Tyr Ser His Ser
                245                 250                 255

CTA AAC GAG GCT GGA AAC GTA TCA ACC TTG GTT AGA TAT CAA GGA ATG       816
Leu Asn Glu Ala Gly Asn Val Ser Thr Leu Val Arg Tyr Gln Gly Met
            260                 265                 270

ATT CAC GGC TTC CTG TCC TTC TAC GAG TGG ATA ACT GCC GGT AAA CTA       864
Ile His Gly Phe Leu Ser Phe Tyr Glu Trp Ile Thr Ala Gly Lys Leu
        275                 280                 285

GCC ATT CAC CAC ATT GCT GGG GTT CTG AGA TCT GTC CTT TA                905
Ala Ile His His Ile Ala Gly Val Leu Arg Ser Val Leu
    290                 295                 300
```

(2) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 978 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:

(A) NAME/KEY: Coding Sequence
(B) LOCATION: 1...975

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

```
GTG GCC TTC TTC GAT ATG CCC CTT GAG GAA CTG AAA AAG TAC CGG CCT        48
Val Ala Phe Phe Asp Met Pro Leu Glu Glu Leu Lys Lys Tyr Arg Pro
 1               5                  10                  15

GAA AGG TAC GAG GAG AAA GAT TTC GAT GAG TTC TGG AGG GAA ACA CTT        96
Glu Arg Tyr Glu Glu Lys Asp Phe Asp Glu Phe Trp Arg Glu Thr Leu
             20                  25                  30

AAA GAA AGC GAA GGA TTC CCT CTG GAT CCC GTC TTT GAA AAG GTG GAC       144
Lys Glu Ser Glu Gly Phe Pro Leu Asp Pro Val Phe Glu Lys Val Asp
         35                  40                  45

TTT CAT CTC AAA ACG GTT GAA ACG TAC GAT GTT ACT TTC TCT GGA TAC       192
Phe His Leu Lys Thr Val Glu Thr Tyr Asp Val Thr Phe Ser Gly Tyr
     50                  55                  60
```

```
AGG GGG CAG AGA ATA AAG GGC TGG CTT CTT GTT CCG AAG TTG GCG GAA       240
Arg Gly Gln Arg Ile Lys Gly Trp Leu Leu Val Pro Lys Leu Ala Glu
65                  70                  75                  80

GAA AAG CTT CCA TGC GTC GTG CAG TAC ATA GGT TAC AAT GGT GGA AGG       288
Glu Lys Leu Pro Cys Val Val Gln Tyr Ile Gly Tyr Asn Gly Gly Arg
                85                  90                  95

GGT TTT CCA CAC GAC TGG CTG TTC TGG CCG TCA ATG GGT TAC ATC TGT       336
Gly Phe Pro His Asp Trp Leu Phe Trp Pro Ser Met Gly Tyr Ile Cys
            100                 105                 110

TTT GTC ATG GAC ACC AGG GGG CAG GGA AGC GGC TGG ATG AAG GGA GAC       384
Phe Val Met Asp Thr Arg Gly Gln Gly Ser Gly Trp Met Lys Gly Asp
        115                 120                 125
ACA CCG GAT TAC CCT GAG GGT CCA GTC GAT CCA CAG TAC CCC GGA TTC       432
Thr Pro Asp Tyr Pro Glu Gly Pro Val Asp Pro Gln Tyr Pro Gly Phe
    130                 135                 140

ATG ACG AGG GGC ATT CTG GAT CCG GGA ACC TAT TAC TAC AGG CGA GTC       480
Met Thr Arg Gly Ile Leu Asp Pro Gly Thr Tyr Tyr Tyr Arg Arg Val
145                 150                 155                 160

TTC GTG GAT GCG GTC AGG GCG GTG GAA GCA GCC ATT TCC TTC CCG AGA       528
Phe Val Asp Ala Val Arg Ala Val Glu Ala Ala Ile Ser Phe Pro Arg
                165                 170                 175

GTG GAT TCC AGG AAG GTG GTG GTG GCC GGA GGC AGT CAG GGT GGG GGA       576
Val Asp Ser Arg Lys Val Val Val Ala Gly Gly Ser Gln Gly Gly Gly
            180                 185                 190

ATC CCC CTT GCG GTG AGT GCC CTG TCG AAC AGG GTG AAG GCT CTG CTC       624
Ile Pro Leu Ala Val Ser Ala Leu Ser Asn Arg Val Lys Ala Leu Leu
        195                 200                 205

TGC GAT GTG CCG TTT CTG TGC CAC TTC AGA AGG GCC GTG CAA CTT GTC       672
Cys Asp Val Pro Phe Leu Cys His Phe Arg Arg Ala Val Gln Leu Val
    210                 215                 220

GAC ACA CAC CCA TAC GTG GAG ATC ACC AAC TTC CTC AAA ACC CAC AGG       720
Asp Thr His Pro Tyr Val Glu Ile Thr Asn Phe Leu Lys Thr His Arg
225                 230                 235                 240

GAC AAA GAG GAG ATT GTT TTC AGA ACA CTT TCC TAC TTC GAT GGT GTG       768
Asp Lys Glu Glu Ile Val Phe Arg Thr Leu Ser Tyr Phe Asp Gly Val
                245                 250                 255

AAC TTT GCA GCA AGG GCA AAG GTG CCC GCC CTG TTT TCC GTT GGG CTC       816
Asn Phe Ala Ala Arg Ala Lys Val Pro Ala Leu Phe Ser Val Gly Leu
            260                 265                 270

ATG GAC ACC ATC TGT CCT CCC TCG ACG GTC TTC GCC GCT TAC AAC CAC       864
Met Asp Thr Ile Cys Pro Pro Ser Thr Val Phe Ala Ala Tyr Asn His
        275                 280                 285
```

```
TAC GCC GGT CCA AAG GAG ATC AGA ATC TAT CCG TAC AAC AAC CAC GAA      912
Tyr Ala Gly Pro Lys Glu Ile Arg Ile Tyr Pro Tyr Asn Asn His Glu
    290                 295                 300

GGT GGA GGT TCT TTC CAG GCA ATT GAG CAG GTG AAA TTC TTG AAG AGA      960
Gly Gly Gly Ser Phe Gln Ala Ile Glu Gln Val Lys Phe Leu Lys Arg
305                 310                 315                 320

CTA TTT GAG GAA GGC TAG                                              978
Leu Phe Glu Glu Gly
                325
```

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 897 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:

(A) NAME/KEY: Coding Sequence
(B) LOCATION: 1...894

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

```
ATG CGC ACC CTC TCC TTC GGT CCG ATG ACC ACA GGG GGA AGC ATT CAC       48
Met Arg Thr Leu Ser Phe Gly Pro Met Thr Thr Gly Gly Ser Ile His
 1               5                  10                  15

ATG GCG ACC ATG GAC GTG ATG CGC GGG CCG GGG ATG CAG CGG CTG TCA       96
Met Ala Thr Met Asp Val Met Arg Gly Pro Gly Met Gln Arg Leu Ser
            20                  25                  30

CAG GGC GCC AGG GAG GCC GCG AAC CAC CCC TGG GCG AAG CGA CTG GGC      144
Gln Gly Ala Arg Glu Ala Ala Asn His Pro Trp Ala Lys Arg Leu Gly
        35                  40                  45

CGC ATG GGC TAC GCG GCC AAG GGC GCC GTG TAC GCC ATC ATC GGC GTG      192
Arg Met Gly Tyr Ala Ala Lys Gly Ala Val Tyr Ala Ile Ile Gly Val
    50                  55                  60

CTC GCG CTG AAG CTC GCG GCG GGC GAG GGC GGC CGG ACC ACG GAC AGC      240
Leu Ala Leu Lys Leu Ala Ala Gly Glu Gly Gly Arg Thr Thr Asp Ser
65                  70                  75                  80

CAC GGC GCG GTG AAC ACC GTG GCG CAC GGG CCC TTC GGC GTC GCG CTG      288
His Gly Ala Val Asn Thr Val Ala His Gly Pro Phe Gly Val Ala Leu
                85                  90                  95

CTG GCG GTG CTG GTG GTG GGC CTG CTG GGC TAC GTG GTC TGG AGG TTC      336
Leu Ala Val Leu Val Val Gly Leu Leu Gly Tyr Val Val Trp Arg Phe
            100                 105                 110
```

```
GCC CAG GCC TTC GTG GAC ACG GAG GAC AAG GGC TCC GAC GCG AAG GGA     384
Ala Gln Ala Phe Val Asp Thr Glu Asp Lys Gly Ser Asp Ala Lys Gly
        115                 120                 125

ATC GCC ACG CGC GCC ATG TAC TTC CTC AGC GGC TGC ATC TAC GCG TCG     432
Ile Ala Thr Arg Ala Met Tyr Phe Leu Ser Gly Cys Ile Tyr Ala Ser
    130                 135                 140

CTG GCC TTC TTC GCC GCG CAG TCC CTG GTG GGC GCC GCG CAC GGC CGG     480
Leu Ala Phe Phe Ala Ala Gln Ser Leu Val Gly Ala Ala His Gly Arg
145                 150                 155                 160

AGC AAG GGG ACG CAG GGC TGG ACG GCC ACG CTG ATG GAG CAG CCC TTT     528
Ser Lys Gly Thr Gln Gly Trp Thr Ala Thr Leu Met Glu Gln Pro Phe
                165                 170                 175

GGC CGC GTG CTG GTG GCG CTG GTG GGG CTG GGC ATC GTG GGC TTC GCG     576
Gly Arg Val Leu Val Ala Leu Val Gly Leu Gly Ile Val Gly Phe Ala
            180                 185                 190

CTG AAG CAG TTC CAC ACC GCG TGG AAG GCG AAG TTC CGG GAG AAG CTC     624
Leu Lys Gln Phe His Thr Ala Trp Lys Ala Lys Phe Arg Glu Lys Leu
        195                 200                 205

ACC CTC ACC GGA CTG GCT GCC CGG AAG CAG CAC CAC ATC GAG CGC ATG     672
Thr Leu Thr Gly Leu Ala Ala Arg Lys Gln His His Ile Glu Arg Met
    210                 215                 220

TGC CAG TTC GGC ATC GCC GCG CGC GGC GTG GTG TTC GCC GTC ATC GGC     720
Cys Gln Phe Gly Ile Ala Ala Arg Gly Val Val Phe Ala Val Ile Gly
225                 230                 235                 240

GGC TTC CTC GTC CGC TCC GCC GTG GAC GCG AAC CCC GGC GAG GCC AAG     768
Gly Phe Leu Val Arg Ser Ala Val Asp Ala Asn Pro Gly Glu Ala Lys
                245                 250                 255

GGC CTG GGA GAG GCC CTG GCC GTC GTC GCG AGG CAG CCG TCC GGC GAC     816
Gly Leu Gly Glu Ala Leu Ala Val Val Ala Arg Gln Pro Ser Gly Asp
            260                 265                 270

GTG CTC CTG GGG GTG GTG GCG GCG GGC CTG GTG GCC TAC GCC GCC TAC     864
Val Leu Leu Gly Val Val Ala Ala Gly Leu Val Ala Tyr Ala Ala Tyr
        275                 280                 285

CTG TTC CTC CAG GCG CGC TAC CGC GAA CTC TAG                         897
Leu Phe Leu Gln Ala Arg Tyr Arg Glu Leu
    290                 295
```

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 914 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:

(A) NAME/KEY: Coding Sequence
(B) LOCATION: 1...912

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

```
ATG AGC AAA TTC GCA ATA CTC TGG GCG TTG ATA ACG GCA TAC CTG CCG          48
Met Ser Lys Phe Ala Ile Leu Trp Ala Leu Ile Thr Ala Tyr Leu Pro
 1               5                   10                  15

GAA CCT GTG ATG AAA CTG GTA TAT TTA GGG CGG CGC GAA ACG CTT GGG          96
Glu Pro Val Met Lys Leu Val Tyr Leu Gly Arg Arg Glu Thr Leu Gly
                 20              25                  30

GCA CGG ACG CTT GAC GTT AAA GCC CAA GCT GTC GGG CGG CTG GCC AAT         144
Ala Arg Thr Leu Asp Val Lys Ala Gln Ala Val Gly Arg Leu Ala Asn
         35                  40                  45

GCA ACA AGA CCT GTC GGG GTG ATT CCG ACG GTC GAG GAA AGC CGG AAG         192
Ala Thr Arg Pro Val Gly Val Ile Pro Thr Val Glu Glu Ser Arg Lys
     50                  55                  60

ATG ACG GAT AAA GCC GTT AGC CTT TTT GAT CAG CCC GCC CCC GAA TTA         240
Met Thr Asp Lys Ala Val Ser Leu Phe Asp Gln Pro Ala Pro Glu Leu
 65                  70                  75                  80

TTC CGT AAA AAA GAC ATT CAG ATT GAC GGG GCT GAA GGG CCT ATT GAT         288
Phe Arg Lys Lys Asp Ile Gln Ile Asp Gly Ala Glu Gly Pro Ile Asp
                 85                  90                  95

GCC CGT ATT TAC AGC GGC CCT GCA AAA CAT CGC CCR CGR CCA ATW CTA         336
Ala Arg Ile Tyr Ser Gly Pro Ala Lys His Arg Xaa Xaa Pro Xaa Leu
            100                 105                 110

GTG TAT TTT CAC GGC GGT GGC TGG GTT CAG GGC AAT CTG GAC AGC CAT         384
Val Tyr Phe His Gly Gly Gly Trp Val Gln Gly Asn Leu Asp Ser His
        115                 120                 125

GAC GGG GTT TGC GGC AAG CTG GCA AAA TGG GCG AAC TGC ATT GTT ATC         432
Asp Gly Val Cys Gly Lys Leu Ala Lys Trp Ala Asn Cys Ile Val Ile
    130                 135                 140

TCG GTC GAT TAT CGT CTA GCG CCC GAA CAC AAA TTT CCT TGT GCG CCG         480
Ser Val Asp Tyr Arg Leu Ala Pro Glu His Lys Phe Pro Cys Ala Pro
145                 150                 155                 160

CTT GAT GCG ATT GCG GCC TAT AAA TGG GTG CGC GCC AAC GCA ACA AAC         528
Leu Asp Ala Ile Ala Ala Tyr Lys Trp Val Arg Ala Asn Ala Thr Asn
                165                 170                 175

CTT GGC GGC GAT CCT GAA CGT ATC GGC GTT GGC GGC GAT AGC GCA GGG         576
```

```
Leu Gly Gly Asp Pro Glu Arg Ile Gly Val Gly Gly Asp Ser Ala Gly
            180                 185                 190

GGC AAT CTT GCC GCC GTT GTC TGC CAA CAA ACC GCC ATG AAC GGC GAG        624
Gly Asn Leu Ala Ala Val Val Cys Gln Gln Thr Ala Met Asn Gly Glu
        195                 200                 205

CGC ACA CCA GAT CTG CAA GTC CTG ATC TAT CCG GCG CTG GAT GCA CGC        672
Arg Thr Pro Asp Leu Gln Val Leu Ile Tyr Pro Ala Leu Asp Ala Arg
    210                 215                 220

ATG ATC TCG ACC TCG ATG GAG GAA TTG CGT GAT GCC TAC ATC TTG CCG        720
Met Ile Ser Thr Ser Met Glu Glu Leu Arg Asp Ala Tyr Ile Leu Pro
225                 230                 235                 240

AAA TCC AGA ATG GAG TAT TTC CTC GGC CTA TAT ACG CGT GGC CCT GAC        768
Lys Ser Arg Met Glu Tyr Phe Leu Gly Leu Tyr Thr Arg Gly Pro Asp
                245                 250                 255

GAT ATC GAG GAC CTT AGG ATG TCG CCA ATT CTC AGG GAT ACC GTC GCG        816
Asp Ile Glu Asp Leu Arg Met Ser Pro Ile Leu Arg Asp Thr Val Ala
            260                 265                 270

GAT CAA CCC CAA GCC TGC ATT GTC ACC TGT GGG TTT GAC CCT GCG CGA        864
Asp Gln Pro Gln Ala Cys Ile Val Thr Cys Gly Phe Asp Pro Ala Arg
        275                 280                 285

CGA CGG GAA CAC CTA CGC CGA ACG CTT AAT TGC CGA GGG GAT AGA CGT        912
Arg Arg Glu His Leu Arg Arg Thr Leu Asn Cys Arg Gly Asp Arg Arg
    290                 295                 300

TA                                                                     914
```

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 926 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:

(A) NAME/KEY: Coding Sequence
(B) LOCATION: 1...924

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

```
GTG AGC ATT CGT CTG CGA CTG TTA AAC TGG TTT TTG AAT ACC TTT GAA         48
Val Ser Ile Arg Leu Arg Leu Leu Asn Trp Phe Leu Asn Thr Phe Glu
  1               5                  10                  15

AAA CCA AAA CTG GCC GCG GCC AAA ACG CCG GAT GAT TTG CGA AAA TCG         96
Lys Pro Lys Leu Ala Ala Ala Lys Thr Pro Asp Asp Leu Arg Lys Ser
```

```
            20                  25                  30

TTT GAA TTA AAG GCG AGG TTT TTG TTT CCG GCG CCA CGT AAA ACA AGG      144
Phe Glu Leu Lys Ala Arg Phe Leu Phe Pro Ala Pro Arg Lys Thr Arg
        35                  40                  45

TTT AGT CAT GAT GTA TTG CAG TCA GGC ATC GGG TCG GTA AAT GCC CAG      192
Phe Ser His Asp Val Leu Gln Ser Gly Ile Gly Ser Val Asn Ala Gln
    50                  55                  60

TGG GCG AAA TCC AAA TCT GCA TCT GAT GAC AGG GTA ATC CTG TAT TTT      240
Trp Ala Lys Ser Lys Ser Ala Ser Asp Asp Arg Val Ile Leu Tyr Phe
65                  70                  75                  80

CAT GGG GGA GGG TAT GTT TTT GGG TCA CCA AAA ACG CAC CGT GCA ATG      288
His Gly Gly Gly Tyr Val Phe Gly Ser Pro Lys Thr His Arg Ala Met
                85                  90                  95

TTG GCG CGC TTG TCG GCA ATG ACA GGT CTT TCT GCG TGC CTT CCA GAT      336
Leu Ala Arg Leu Ser Ala Met Thr Gly Leu Ser Ala Cys Leu Pro Asp
            100                 105                 110

TAT AGG TTG GCA CCA GAG CAC CCA TTT CCA GCC GCG ATC GAA GAT GCA      384
Tyr Arg Leu Ala Pro Glu His Pro Phe Pro Ala Ala Ile Glu Asp Ala
        115                 120                 125

GTT TTA TCG TAT AAA TGT TTA CTA GAG CGA GCA ATC GAG CCC CAA AAT      432
Val Leu Ser Tyr Lys Cys Leu Leu Glu Arg Ala Ile Glu Pro Gln Asn
    130                 135                 140

ATT ATA CTG GGG GGG GAC AGT GCT GGT GGC GGT TTG GTT CTT GCT TTG      480
Ile Ile Leu Gly Gly Asp Ser Ala Gly Gly Gly Leu Val Leu Ala Leu
145                 150                 155                 160

CTT GCA GAA ATC AAG GCC CAA TCC TTG CCC AAA CCT GCT GGC GTT TTT      528
Leu Ala Glu Ile Lys Ala Gln Ser Leu Pro Lys Pro Ala Gly Val Phe
                165                 170                 175

GCC TTG TCG CCT TTG GTT GAT TTA TCA TTT TCG GGC CTT TCG TTT TCT      576
Ala Leu Ser Pro Leu Val Asp Leu Ser Phe Ser Gly Leu Ser Phe Ser
            180                 185                 190

AAA AAT GCC CAA ACC GAT GTG ATG TTG CCC GCA TCA CGG GCT GCG GAT      624
Lys Asn Ala Gln Thr Asp Val Met Leu Pro Ala Ser Arg Ala Ala Asp
        195                 200                 205

ATG GCG ACC TTG TAT TTG GAT GGG GCC GAT GCA GAT GAT CCA CGT GCA      672
Met Ala Thr Leu Tyr Leu Asp Gly Ala Asp Ala Asp Asp Pro Arg Ala
    210                 215                 220

TCG CCG CTG CAG GCG GAT TTT TCT GGC ATG CCG CCT GTA TTT CTG ACA      720
Ser Pro Leu Gln Ala Asp Phe Ser Gly Met Pro Pro Val Phe Leu Thr
225                 230                 235                 240

GCA AGT GAC AGT GAA ATC CTG TTG GAT GAT TGC CTG CGG ATG GCG GAT      768
```

```
Ala Ser Asp Ser Glu Ile Leu Leu Asp Asp Cys Leu Arg Met Ala Asp
                245                 250                 255

CAC TTG CGT GCG CAA GGT GTC GTT GTG ACA GAC CGG ATT GTT GAA AAC      816
His Leu Arg Ala Gln Gly Val Val Val Thr Asp Arg Ile Val Glu Asn
            260                 265                 270

CAT CCA CAT GTT TGG CAT ATT TTT CAA CGC CTT CTA CCC GAA GCA GAT      864
His Pro His Val Trp His Ile Phe Gln Arg Leu Leu Pro Glu Ala Asp
        275                 280                 285

CAG GGG CTG CGG GCG ATT GCC GCG TGG ATT AAA CCT CTT TTA TCA GGT      912
Gln Gly Leu Arg Ala Ile Ala Ala Trp Ile Lys Pro Leu Leu Ser Gly
    290                 295                 300

TCA AAC GAA AGC TA                                                   926
Ser Asn Glu Ser
305
```

(2) INFORMATION FOR SEQ ID NO:50:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 713 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:

(A) NAME/KEY: Coding Sequence
(B) LOCATION: 1...711

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

```
ATG CTT ACA TTT AAT GTT TTA TAT GGT ATG ATG AAA CAA AAA CTA GCA       48
Met Leu Thr Phe Asn Val Leu Tyr Gly Met Met Lys Gln Lys Leu Ala
 1               5                  10                  15

GCA ATT CTC ATG TTT TTA GGG CTA TCA GCA GCA GAG GCT CAA GAC TGG       96
Ala Ile Leu Met Phe Leu Gly Leu Ser Ala Ala Glu Ala Gln Asp Trp
            20                  25                  30

CCT GAC CTA CAG AAA TAT CGT AGT GCT AAT AAA GAA GCC AAA TTA CTT      144
Pro Asp Leu Gln Lys Tyr Arg Ser Ala Asn Lys Glu Ala Lys Leu Leu
        35                  40                  45

CCA AAG GAA AAC CGG AAG GTG GTT TTT ATG GGC AAC TCC ATT ACA GAA      192
Pro Lys Glu Asn Arg Lys Val Val Phe Met Gly Asn Ser Ile Thr Glu
    50                  55                  60

GCC TGG ATT AGT CAG CGA CCT GAG TTT TTT AGT GAA AAT GGG TTT ATC      240
Ala Trp Ile Ser Gln Arg Pro Glu Phe Phe Ser Glu Asn Gly Phe Ile
 65                 70                  75                  80
```

```
GGT CGA GGC ATC AGT GGC CAG ACA ACC CCT CAG ATG TTG TTG AGA TTC       288
Gly Arg Gly Ile Ser Gly Gln Thr Thr Pro Gln Met Leu Leu Arg Phe
                 85                  90                  95

CGA CAG GAT GTG ATA GAC CTG CAG CCA AAG GCT GTA GTG ATA CTA GCT       336
Arg Gln Asp Val Ile Asp Leu Gln Pro Lys Ala Val Val Ile Leu Ala
            100                 105                 110
GGT ACC AAT GAC GTA GCT CAA AAT ACC GGG CCG ATG ACC ATT GAG GAA       384
Gly Thr Asn Asp Val Ala Gln Asn Thr Gly Pro Met Thr Ile Glu Glu
        115                 120                 125

TCG CTT GCT AAC ATT AAG TCT ATG GTG GAG CTG GCG CAA GCC AAT GGG       432
Ser Leu Ala Asn Ile Lys Ser Met Val Glu Leu Ala Gln Ala Asn Gly
    130                 135                 140

ATC ACG CCT GTT TTG TGT ACC GTG CTG CCT GCA GAT CGT TTC AGC TGG       480
Ile Thr Pro Val Leu Cys Thr Val Leu Pro Ala Asp Arg Phe Ser Trp
145                 150                 155                 160

CGA CCT GAG CTT ACA CCC GCA GAA ACT ATC ATT GCC CTC AAT CAG CTC       528
Arg Pro Glu Leu Thr Pro Ala Glu Thr Ile Ile Ala Leu Asn Gln Leu
                165                 170                 175

ATT AAG CAA TAT GCC GAG GCA CAG GGC CTG GCC CTG GTG GAT TAT CAT       576
Ile Lys Gln Tyr Ala Glu Ala Gln Gly Leu Ala Leu Val Asp Tyr His
            180                 185                 190

GCT GCA CTC ACC AAT AAA GGT GGA GGA CTT CCG GTG AAA TAC GGA GAA       624
Ala Ala Leu Thr Asn Lys Gly Gly Gly Leu Pro Val Lys Tyr Gly Glu
        195                 200                 205

GAT GGT GTG CAT CCA AAT GTA GCA GGC TAT CAG GTG ATG GAA AAC ATT       672
Asp Gly Val His Pro Asn Val Ala Gly Tyr Gln Val Met Glu Asn Ile
    210                 215                 220

GTT TTA CCG GTC ATT TCC AGC GAG TTG GCA AAG CTG AAG TA                713
Val Leu Pro Val Ile Ser Ser Glu Leu Ala Lys Leu Lys
225                 230                 235
```

(2) INFORMATION FOR SEQ ID NO:51:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 978 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:

(A) NAME/KEY: Coding Sequence
(B) LOCATION: 1...975

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

```
ATG GCC TTC TTC GAT TTA CCA CTC GAA GAA CTG AAG AAA TAT CGT CCA          48
Met Ala Phe Phe Asp Leu Pro Leu Glu Glu Leu Lys Lys Tyr Arg Pro
 1               5                   10                  15

GAG CGG TAC GAA GAG AAA GAC TTC GAT GAG TTC TGG GAA GAG ACA CTC          96
Glu Arg Tyr Glu Glu Lys Asp Phe Asp Glu Phe Trp Glu Glu Thr Leu
             20                  25                  30

GCA GAG AGC GAA AAG TTC CCC TTA GAC CCC GTC TTC GAG AGG ATG GAG         144
Ala Glu Ser Glu Lys Phe Pro Leu Asp Pro Val Phe Glu Arg Met Glu
         35                  40                  45

TCT CAC CTC AAA ACA GTC GAA GCG TAC GAT GTC ACC TTC TCC GGA TAC         192
Ser His Leu Lys Thr Val Glu Ala Tyr Asp Val Thr Phe Ser Gly Tyr
     50                  55                  60

AGG GGA CAG AGG ATC AAA GGG TGG CTC CTT GTT CCA AAA CTG GAA GAA         240
Arg Gly Gln Arg Ile Lys Gly Trp Leu Leu Val Pro Lys Leu Glu Glu
 65                  70                  75                  80

GAA AAA CTT CCC TGC GTT GTG CAG TAC ATA GGA TAC AAC GGT GGA AGA         288
Glu Lys Leu Pro Cys Val Val Gln Tyr Ile Gly Tyr Asn Gly Gly Arg
                 85                  90                  95

GGA TTC CCT CAC GAC TGG CTG TTC TGG CCT TCT ATG GGT TAC ATA TGT         336
Gly Phe Pro His Asp Trp Leu Phe Trp Pro Ser Met Gly Tyr Ile Cys
            100                 105                 110

TTC GTC ATG GAT ACT CGA GGT CAG GGA AGC GGC TGG CTG AAA GGA GAC         384
Phe Val Met Asp Thr Arg Gly Gln Gly Ser Gly Trp Leu Lys Gly Asp
        115                 120                 125

ACA CCG GAT TAC CCT GAG GGT CCC GTT GAC CCT CAG TAT CCA GGA TTC         432
Thr Pro Asp Tyr Pro Glu Gly Pro Val Asp Pro Gln Tyr Pro Gly Phe
    130                 135                 140

ATG ACA AGA GGA ATA CTG GAT CCC AGA ACT TAC TAC TAC AGA CGA GTC         480
Met Thr Arg Gly Ile Leu Asp Pro Arg Thr Tyr Tyr Tyr Arg Arg Val
145                 150                 155                 160

TTC ACG GAC GCT GTC AGA GCC GTT GAA GCT GCT GCT TCT TTT CCT CAG         528
Phe Thr Asp Ala Val Arg Ala Val Glu Ala Ala Ala Ser Phe Pro Gln
                165                 170                 175

GTA GAT CAA GAA AGA ATC GTG ATA GCT GGA GGC AGT CAG GGT GGC GGA         576
Val Asp Gln Glu Arg Ile Val Ile Ala Gly Gly Ser Gln Gly Gly Gly
            180                 185                 190

ATA GCC CTT GCG GTG AGC GCT CTC TCA AAG AAA GCA AAG GCT CTT CTG         624
Ile Ala Leu Ala Val Ser Ala Leu Ser Lys Lys Ala Lys Ala Leu Leu
        195                 200                 205

TGC GAT GTG CCG TTT CTG TGT CAC TTC AGA AGA GCA GTA CAG CTT GTG         672
```

```
Cys Asp Val Pro Phe Leu Cys His Phe Arg Arg Ala Val Gln Leu Val
    210                 215                 220

GAT ACG CAT CCA TAC GCG GAG ATC ACG AAC TTT CTA AAG ACC CAC AGA      720
Asp Thr His Pro Tyr Ala Glu Ile Thr Asn Phe Leu Lys Thr His Arg
225                 230                 235                 240
GAC AAG GAA GAA ATC GTG TTC AGG ACT CTT TCC TAT TTC GAT GGA GTG      768
Asp Lys Glu Glu Ile Val Phe Arg Thr Leu Ser Tyr Phe Asp Gly Val
                245                 250                 255

AAC TTC GCA GCC AGA GCG AAG ATC CCT GCG CTG TTT TCT GTG GGT CTC      816
Asn Phe Ala Ala Arg Ala Lys Ile Pro Ala Leu Phe Ser Val Gly Leu
            260                 265                 270

ATG GAC AAC ATT TGT CCT CCT TCA ACG GTT TTC GCT GCC TAC AAT TAC      864
Met Asp Asn Ile Cys Pro Pro Ser Thr Val Phe Ala Ala Tyr Asn Tyr
        275                 280                 285

TAC GCT GGA CCG AAG GAA ATC AGA ATC TAT CCG TAC AAC AAC CAC GAG      912
Tyr Ala Gly Pro Lys Glu Ile Arg Ile Tyr Pro Tyr Asn Asn His Glu
    290                 295                 300

GGA GGA GGC TCT TTC CAA GCG GTT GAA CAG GTG AAA TTC TTG AAA AAA      960
Gly Gly Gly Ser Phe Gln Ala Val Glu Gln Val Lys Phe Leu Lys Lys
305                 310                 315                 320

CTA TTT GAG AAA GGC TAA                                              978
Leu Phe Glu Lys Gly
                325
```

(2) INFORMATION FOR SEQ ID NO:52:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 661 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:

(A) NAME/KEY: Coding Sequence
(B) LOCATION: 1...630

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

```
TTG AAG TAC TTC AAA GCC CGG CTT GCC GGC ATC ACC TTG CTC GGC CTG       48
Leu Lys Tyr Phe Lys Ala Arg Leu Ala Gly Ile Thr Leu Leu Gly Leu
 1               5                  10                  15

CTG GCC TGC ACC TCG GCC TCG GCG CAG ACC GAG CCC ATC GTG TTC GTG       96
Leu Ala Cys Thr Ser Ala Ser Ala Gln Thr Glu Pro Ile Val Phe Val
            20                  25                  30
```

```
CAC GGC TAT TCC GGC AGC GCA TCC AAC TGG GAC ACC ATG CTG GGC CGC        144
His Gly Tyr Ser Gly Ser Ala Ser Asn Trp Asp Thr Met Leu Gly Arg
         35                  40                  45

TTC CGG TCG AAC GGT TAT GCG TCC GGC TCG CTC TAC ACC TTC AAC TAC        192
Phe Arg Ser Asn Gly Tyr Ala Ser Gly Ser Leu Tyr Thr Phe Asn Tyr
     50                  55                  60

AAC TCG TTG GTC AGC AGC AAC CGC ACC AGC GCC AGC GAG CTG CGC AGC        240
Asn Ser Leu Val Ser Ser Asn Arg Thr Ser Ala Ser Glu Leu Arg Ser
 65                  70                  75                  80

TTC GTC AAC ACC GTG CGT TCG CGC CAC GGC AAC GCC CGC ATC GCG CTG        288
Phe Val Asn Thr Val Arg Ser Arg His Gly Asn Ala Arg Ile Ala Leu
                 85                  90                  95

GTC GCC CAC TCC AAC GGC GGG CTG GTG TCG CGC TGG TAT CGC GCG GAG        336
Val Ala His Ser Asn Gly Gly Leu Val Ser Arg Trp Tyr Arg Ala Glu
            100                 105                 110

CTG GGC GGC GAA ACG GCC ACC CGC CGC TTC GTG ACG CTG GGC ACG CCG        384
Leu Gly Gly Glu Thr Ala Thr Arg Arg Phe Val Thr Leu Gly Thr Pro
        115                 120                 125

CAC CGG GGC ACC ACC TGG GCC TAT GCG TGC TAC AGC CCC GCA TGT TTC        432
His Arg Gly Thr Thr Trp Ala Tyr Ala Cys Tyr Ser Pro Ala Cys Phe
    130                 135                 140

GAG ATG CGC CCC GGC TCC AGC TTG CTG ACC ACG CTG GGC TCG CGT GCC        480
Glu Met Arg Pro Gly Ser Ser Leu Leu Thr Thr Leu Gly Ser Arg Ala
145                 150                 155                 160

TGC GAC CGC TCG CTG TGG TCG AAC ACC GAC GGC ATC ATC CTG CCG GCG        528
Cys Asp Arg Ser Leu Trp Ser Asn Thr Asp Gly Ile Ile Leu Pro Ala
                165                 170                 175

TCC AGC GCG CAG TGT GGT GTC AGC ACG CGC ACT GCC GAC GTC AGC CAT        576
Ser Ser Ala Gln Cys Gly Val Ser Thr Arg Thr Ala Asp Val Ser His
            180                 185                 190

CTC GAC CTG CTG ACC GAC TCT CGC GTG TAC ACG CAG TTG CGC ACG CAG        624
Leu Asp Leu Leu Thr Asp Ser Arg Val Tyr Thr Gln Leu Arg Thr Gln
        195                 200                 205

TTG CAA TGAGGGTGAC GGTGCACCGA ACGTGCACCT G                             661
Leu Gln
    210
```

(2) INFORMATION FOR SEQ ID NO:53:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 201 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

```
Met Lys Val Lys His Val Ile Val Leu His Gly Leu Tyr Met Ser Gly
  1               5                  10                  15
Leu Val Met Arg Pro Leu Cys Ser Arg Leu Glu Glu Ser Gly Val Lys
             20                  25                  30
Val Leu Asn Leu Thr Tyr Asn Thr Arg Asp Pro Asn Arg Asp Ala Ile
         35                  40                  45
Phe Thr Gln Ile Asp Glu Phe Ile Ser Asn Glu Pro Ser Ala Leu Val
     50                  55                  60
Cys His Ser Met Gly Gly Leu Val Ala Arg Ala Tyr Leu Glu Ala Asn
 65                  70                  75                  80
Ser Ala Pro Ser His His Val Glu Lys Val Ile Thr Leu Gly Thr Pro
                 85                  90                  95
His Thr Gly Ser His Ile Ala Glu Lys Met Gln Gln Lys Gly Phe Glu
            100                 105                 110
Leu Leu Leu Lys Asn Ser Val Glu Phe Leu Leu Ser Lys Asn Gly Asp
        115                 120                 125
Trp Pro Phe Lys Ala Lys Leu Tyr Ser Ile Ala Gly Asp Leu Pro Ile
    130                 135                 140
Gly Leu Met Pro Leu Ile Val Lys Gly Ser Arg Ser Asp Gly Thr Val
145                 150                 155                 160
Leu Leu Asp Glu Thr Lys Leu Lys Gly Met Ala Glu His Lys Val Phe
                165                 170                 175
His Leu Ser His Thr Ser Met Ile Tyr Ser Arg Gln Val Val Asn Tyr
            180                 185                 190
Ile Leu Glu Arg Leu Asn Glu Asp Ile
        195                 200
```

(2) INFORMATION FOR SEQ ID NO:54:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 259 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

```
Met Ile Lys Asn Phe Asp Arg Glu Asn Ser Ser Leu Val Leu Ser Gly
  1               5                  10                  15
Gly Gly Ala Leu Gly Ile Ala His Leu Gly Val Leu His Asp Leu Glu
             20                  25                  30
Lys Gln Asn Ile Val Pro Asn Glu Ile Val Gly Thr Ser Met Gly Gly
         35                  40                  45
Ile Ile Gly Ala Ser Met Ala Ile Gly Met Lys Glu Lys Glu Ile Leu
     50                  55                  60
Glu Glu Ile Lys Asn Phe Ser Asn Val Phe Asn Trp Ile Lys Phe Ser
```

```
 65                  70                  75                  80
Phe Ser Gly Asn Ser Val Val Asp Asn Glu Lys Ile Ala Lys Ile Phe
                 85                  90                  95
Asp Thr Leu Phe Lys Asp Arg Lys Met Thr Asp Thr Val Ile Pro Leu
            100                 105                 110
Lys Leu Ile Ala Thr Asn Leu His Asn Gly His Lys Lys Val Phe Thr
        115                 120                 125
Ala Ser Asp Asp Val Leu Ile Lys Asp Ala Ile Leu Ser Thr Met Ala
    130                 135                 140
Ile Pro Gly Val Phe Glu Glu His Ile Ile Asp Gly Glu Thr Tyr Gly
145                 150                 155                 160
Asp Gly Phe Leu Cys Glu Asn Leu Gly Val Asn Glu Ala Thr Phe Asn
                165                 170                 175
Asp Val Leu Ala Val Asp Val Met Gly Glu Asn Ser Phe Glu Lys Ala
            180                 185                 190
Met Pro Asp Asn Phe Phe Lys Thr Ser Asn Val Leu Glu Met Phe Glu
        195                 200                 205
Lys Ser Met Arg Leu Phe Ile Tyr Asn Gln Thr Gln Thr His Ile Lys
    210                 215                 220
Asn Ala Asn Lys Asn Ile Tyr Leu Ile Glu Pro Val Thr Lys Glu Tyr
225                 230                 235                 240
Lys Thr Phe Gln Phe His Lys His Lys Glu Ile Arg Ala Leu Gly Leu
                245                 250                 255
Gly Leu Leu
```

(2) INFORMATION FOR SEQ ID NO:55:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 301 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

```
Met Pro Leu His Pro Lys Val Lys Lys Leu Leu Ser Gln Leu Pro Pro
 1               5                   10                  15
Gln Asp Phe Ser Arg Asn Val Gln Asp Leu Arg Lys Ala Trp Asp Leu
             20                  25                  30
Pro Phe Ser Gly Arg Arg Glu Thr Leu Lys Arg Val Glu Asp Leu Glu
         35                  40                  45
Ile Pro Thr Arg Asp Ala Arg Ile Arg Ala Arg Val Tyr Thr Pro Ser
     50                  55                  60
Ser Lys Glu Asn Leu Pro Val Leu Val Tyr Tyr His Gly Gly Gly Phe
 65                  70                  75                  80
Val Phe Gly Ser Val Asp Ser Tyr Asp Gly Leu Ala Ser Leu Ile Ala
                 85                  90                  95
Lys Glu Ser Gly Ile Ala Val Ile Ser Val Glu Tyr Arg Leu Ala Pro
            100                 105                 110
Glu His Lys Phe Pro Thr Ala Val Asn Asp Ser Trp Asp Ala Leu Leu
```

```
        115                 120                 125
Trp Ile Ala Glu Asn Gly Gly Lys Leu Gly Leu Asp Thr Ser Arg Leu
    130                 135                 140
Ala Val Ala Gly Asp Ser Ala Gly Gly Asn Leu Ser Ala Val Val Ser
145                 150                 155                 160
Leu Leu Asp Arg Asp Gln Gly Lys Gly Leu Val Ser Tyr Gln Val Leu
                165                 170                 175
Ile Tyr Pro Ala Val Asn Met Val Asp Asn Ser Pro Ser Val Arg Glu
            180                 185                 190
Tyr Gly Glu Gly Tyr Phe Leu Thr Arg Ser Met Met Asn Trp Phe Gly
        195                 200                 205
Thr Met Tyr Phe Ser Ser Gly Arg Glu Ala Val Ser Pro Tyr Ala Ser
    210                 215                 220
Pro Ala Leu Ala Asp Leu His Asn Leu Pro Pro Ser Leu Val Ile Thr
225                 230                 235                 240
Ala Glu Tyr Asp Pro Leu Arg Asp Gln Gly Glu Thr Tyr Ser His Ser
                245                 250                 255
Leu Asn Glu Ala Gly Asn Val Ser Thr Leu Val Arg Tyr Gln Gly Met
            260                 265                 270
Ile His Gly Phe Leu Ser Phe Tyr Glu Trp Ile Thr Ala Gly Lys Leu
        275                 280                 285
Ala Ile His His Ile Ala Gly Val Leu Arg Ser Val Leu
    290                 295                 300
```

(2) INFORMATION FOR SEQ ID NO:56:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 325 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

```
Val Ala Phe Phe Asp Met Pro Leu Glu Glu Leu Lys Lys Tyr Arg Pro
1               5                   10                  15
Glu Arg Tyr Glu Glu Lys Asp Phe Asp Glu Phe Trp Arg Glu Thr Leu
            20                  25                  30
Lys Glu Ser Glu Gly Phe Pro Leu Asp Pro Val Phe Glu Lys Val Asp
        35                  40                  45
Phe His Leu Lys Thr Val Glu Thr Tyr Asp Val Thr Phe Ser Gly Tyr
    50                  55                  60
Arg Gly Gln Arg Ile Lys Gly Trp Leu Leu Val Pro Lys Leu Ala Glu
65                  70                  75                  80
Glu Lys Leu Pro Cys Val Val Gln Tyr Ile Gly Tyr Asn Gly Gly Arg
                85                  90                  95
Gly Phe Pro His Asp Trp Leu Phe Trp Pro Ser Met Gly Tyr Ile Cys
            100                 105                 110
Phe Val Met Asp Thr Arg Gly Gln Gly Ser Gly Trp Met Lys Gly Asp
        115                 120                 125
Thr Pro Asp Tyr Pro Glu Gly Pro Val Asp Pro Gln Tyr Pro Gly Phe
```

```
    130                 135                 140
Met Thr Arg Gly Ile Leu Asp Pro Gly Thr Tyr Tyr Tyr Arg Arg Val
145                 150                 155                 160
Phe Val Asp Ala Val Arg Ala Val Glu Ala Ala Ile Ser Phe Pro Arg
                165                 170                 175
Val Asp Ser Arg Lys Val Val Val Ala Gly Gly Ser Gln Gly Gly Gly
            180                 185                 190
Ile Pro Leu Ala Val Ser Ala Leu Ser Asn Arg Val Lys Ala Leu Leu
        195                 200                 205
Cys Asp Val Pro Phe Leu Cys His Phe Arg Arg Ala Val Gln Leu Val
    210                 215                 220
Asp Thr His Pro Tyr Val Glu Ile Thr Asn Phe Leu Lys Thr His Arg
225                 230                 235                 240
Asp Lys Glu Glu Ile Val Phe Arg Thr Leu Ser Tyr Phe Asp Gly Val
                245                 250                 255
Asn Phe Ala Ala Arg Ala Lys Val Pro Ala Leu Phe Ser Val Gly Leu
            260                 265                 270
Met Asp Thr Ile Cys Pro Pro Ser Thr Val Phe Ala Ala Tyr Asn His
        275                 280                 285
Tyr Ala Gly Pro Lys Glu Ile Arg Ile Tyr Pro Tyr Asn Asn His Glu
    290                 295                 300
Gly Gly Gly Ser Phe Gln Ala Ile Glu Gln Val Lys Phe Leu Lys Arg
305                 310                 315                 320
Leu Phe Glu Glu Gly
                325
```

(2) INFORMATION FOR SEQ ID NO:57:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 298 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

```
Met Arg Thr Leu Ser Phe Gly Pro Met Thr Thr Gly Gly Ser Ile His
1               5                   10                  15
Met Ala Thr Met Asp Val Met Arg Gly Pro Gly Met Gln Arg Leu Ser
            20                  25                  30
Gln Gly Ala Arg Glu Ala Ala Asn His Pro Trp Ala Lys Arg Leu Gly
        35                  40                  45
Arg Met Gly Tyr Ala Ala Lys Gly Ala Val Tyr Ala Ile Ile Gly Val
    50                  55                  60
Leu Ala Leu Lys Leu Ala Ala Gly Glu Gly Gly Arg Thr Thr Asp Ser
65                  70                  75                  80
His Gly Ala Val Asn Thr Val Ala His Gly Pro Phe Gly Val Ala Leu
                85                  90                  95
Leu Ala Val Leu Val Val Gly Leu Leu Gly Tyr Val Val Trp Arg Phe
            100                 105                 110
Ala Gln Ala Phe Val Asp Thr Glu Asp Lys Gly Ser Asp Ala Lys Gly
        115                 120                 125
```

```
Ile Ala Thr Arg Ala Met Tyr Phe Leu Ser Gly Cys Ile Tyr Ala Ser
    130                 135                 140
Leu Ala Phe Phe Ala Ala Gln Ser Leu Val Gly Ala Ala His Gly Arg
145                 150                 155                 160
Ser Lys Gly Thr Gln Gly Trp Thr Ala Thr Leu Met Glu Gln Pro Phe
                165                 170                 175
Gly Arg Val Leu Val Ala Leu Val Gly Leu Gly Ile Val Gly Phe Ala
            180                 185                 190
Leu Lys Gln Phe His Thr Ala Trp Lys Ala Lys Phe Arg Glu Lys Leu
        195                 200                 205
Thr Leu Thr Gly Leu Ala Ala Arg Lys Gln His His Ile Glu Arg Met
    210                 215                 220
Cys Gln Phe Gly Ile Ala Ala Arg Gly Val Val Phe Ala Val Ile Gly
225                 230                 235                 240
Gly Phe Leu Val Arg Ser Ala Val Asp Ala Asn Pro Gly Glu Ala Lys
                245                 250                 255
Gly Leu Gly Glu Ala Leu Ala Val Val Ala Arg Gln Pro Ser Gly Asp
            260                 265                 270
Val Leu Leu Gly Val Val Ala Ala Gly Leu Val Ala Tyr Ala Ala Tyr
        275                 280                 285
Leu Phe Leu Gln Ala Arg Tyr Arg Glu Leu
    290                 295
```

(2) INFORMATION FOR SEQ ID NO:58:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 304 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

```
Met Ser Lys Phe Ala Ile Leu Trp Ala Leu Ile Thr Ala Tyr Leu Pro
  1               5                  10                  15
Glu Pro Val Met Lys Leu Val Tyr Leu Gly Arg Arg Glu Thr Leu Gly
             20                  25                  30
Ala Arg Thr Leu Asp Val Lys Ala Gln Ala Val Gly Arg Leu Ala Asn
         35                  40                  45
Ala Thr Arg Pro Val Gly Val Ile Pro Thr Val Glu Glu Ser Arg Lys
     50                  55                  60
Met Thr Asp Lys Ala Val Ser Leu Phe Asp Gln Pro Ala Pro Glu Leu
 65                  70                  75                  80
Phe Arg Lys Lys Asp Ile Gln Ile Asp Gly Ala Glu Gly Pro Ile Asp
                 85                  90                  95
Ala Arg Ile Tyr Ser Gly Pro Ala Lys His Arg Xaa Xaa Pro Xaa Leu
            100                 105                 110
Val Tyr Phe His Gly Gly Gly Trp Val Gln Gly Asn Leu Asp Ser His
        115                 120                 125
```

```
Asp Gly Val Cys Gly Lys Leu Ala Lys Trp Ala Asn Cys Ile Val Ile
    130                 135                 140
Ser Val Asp Tyr Arg Leu Ala Pro Glu His Lys Phe Pro Cys Ala Pro
145                 150                 155                 160
Leu Asp Ala Ile Ala Ala Tyr Lys Trp Val Arg Ala Asn Ala Thr Asn
                165                 170                 175
Leu Gly Gly Asp Pro Glu Arg Ile Gly Val Gly Gly Asp Ser Ala Gly
            180                 185                 190
Gly Asn Leu Ala Ala Val Val Cys Gln Gln Thr Ala Met Asn Gly Glu
        195                 200                 205
Arg Thr Pro Asp Leu Gln Val Leu Ile Tyr Pro Ala Leu Asp Ala Arg
    210                 215                 220
Met Ile Ser Thr Ser Met Glu Glu Leu Arg Asp Ala Tyr Ile Leu Pro
225                 230                 235                 240
Lys Ser Arg Met Glu Tyr Phe Leu Gly Leu Tyr Thr Arg Gly Pro Asp
                245                 250                 255
Asp Ile Glu Asp Leu Arg Met Ser Pro Ile Leu Arg Asp Thr Val Ala
            260                 265                 270
Asp Gln Pro Gln Ala Cys Ile Val Thr Cys Gly Phe Asp Pro Ala Arg
        275                 280                 285
Arg Arg Glu His Leu Arg Arg Thr Leu Asn Cys Arg Gly Asp Arg Arg
    290                 295                 300
```

(2) INFORMATION FOR SEQ ID NO:59:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 308 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

```
Val Ser Ile Arg Leu Arg Leu Leu Asn Trp Phe Leu Asn Thr Phe Glu
 1               5                  10                  15
Lys Pro Lys Leu Ala Ala Ala Lys Thr Pro Asp Asp Leu Arg Lys Ser
             20                  25                  30
Phe Glu Leu Lys Ala Arg Phe Leu Phe Pro Ala Pro Arg Lys Thr Arg
         35                  40                  45
Phe Ser His Asp Val Leu Gln Ser Gly Ile Gly Ser Val Asn Ala Gln
     50                  55                  60
Trp Ala Lys Ser Lys Ser Ala Ser Asp Asp Arg Val Ile Leu Tyr Phe
 65                  70                  75                  80
His Gly Gly Gly Tyr Val Phe Gly Ser Pro Lys Thr His Arg Ala Met
                 85                  90                  95
Leu Ala Arg Leu Ser Ala Met Thr Gly Leu Ser Ala Cys Leu Pro Asp
            100                 105                 110
Tyr Arg Leu Ala Pro Glu His Pro Phe Pro Ala Ala Ile Glu Asp Ala
        115                 120                 125
Val Leu Ser Tyr Lys Cys Leu Leu Glu Arg Ala Ile Glu Pro Gln Asn
    130                 135                 140
```

```
Ile Ile Leu Gly Gly Asp Ser Ala Gly Gly Gly Leu Val Leu Ala Leu
145                 150                 155                 160
Leu Ala Glu Ile Lys Ala Gln Ser Leu Pro Lys Pro Ala Gly Val Phe
                165                 170                 175
Ala Leu Ser Pro Leu Val Asp Leu Ser Phe Ser Gly Leu Ser Phe Ser
            180                 185                 190
Lys Asn Ala Gln Thr Asp Val Met Leu Pro Ala Ser Arg Ala Ala Asp
        195                 200                 205
Met Ala Thr Leu Tyr Leu Asp Gly Ala Asp Ala Asp Asp Pro Arg Ala
    210                 215                 220
Ser Pro Leu Gln Ala Asp Phe Ser Gly Met Pro Pro Val Phe Leu Thr
225                 230                 235                 240
Ala Ser Asp Ser Glu Ile Leu Leu Asp Asp Cys Leu Arg Met Ala Asp
                245                 250                 255
His Leu Arg Ala Gln Gly Val Val Val Thr Asp Arg Ile Val Glu Asn
            260                 265                 270
His Pro His Val Trp His Ile Phe Gln Arg Leu Leu Pro Glu Ala Asp
        275                 280                 285
Gln Gly Leu Arg Ala Ile Ala Ala Trp Ile Lys Pro Leu Leu Ser Gly
    290                 295                 300
Ser Asn Glu Ser
305
```

(2) INFORMATION FOR SEQ ID NO:60:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 237 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

```
Met Leu Thr Phe Asn Val Leu Tyr Gly Met Met Lys Gln Lys Leu Ala
 1               5                   10                  15
Ala Ile Leu Met Phe Leu Gly Leu Ser Ala Ala Glu Ala Gln Asp Trp
             20                  25                  30
Pro Asp Leu Gln Lys Tyr Arg Ser Ala Asn Lys Glu Ala Lys Leu Leu
         35                  40                  45
Pro Lys Glu Asn Arg Lys Val Val Phe Met Gly Asn Ser Ile Thr Glu
     50                  55                  60
Ala Trp Ile Ser Gln Arg Pro Glu Phe Phe Ser Glu Asn Gly Phe Ile
 65                  70                  75                  80
Gly Arg Gly Ile Ser Gly Gln Thr Thr Pro Gln Met Leu Leu Arg Phe
                 85                  90                  95
Arg Gln Asp Val Ile Asp Leu Gln Pro Lys Ala Val Val Ile Leu Ala
            100                 105                 110
Gly Thr Asn Asp Val Ala Gln Asn Thr Gly Pro Met Thr Ile Glu Glu
        115                 120                 125
Ser Leu Ala Asn Ile Lys Ser Met Val Glu Leu Ala Gln Ala Asn Gly
    130                 135                 140
```

```
Ile Thr Pro Val Leu Cys Thr Val Leu Pro Ala Asp Arg Phe Ser Trp
145                 150                 155                 160
Arg Pro Glu Leu Thr Pro Ala Glu Thr Ile Ile Ala Leu Asn Gln Leu
                165                 170                 175
Ile Lys Gln Tyr Ala Glu Ala Gln Gly Leu Ala Leu Val Asp Tyr His
            180                 185                 190
Ala Ala Leu Thr Asn Lys Gly Gly Gly Leu Pro Val Lys Tyr Gly Glu
        195                 200                 205
Asp Gly Val His Pro Asn Val Ala Gly Tyr Gln Val Met Glu Asn Ile
    210                 215                 220
Val Leu Pro Val Ile Ser Ser Glu Leu Ala Lys Leu Lys
225                 230                 235
```

(2) INFORMATION FOR SEQ ID NO:61:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 325 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

```
Met Ala Phe Phe Asp Leu Pro Leu Glu Glu Leu Lys Lys Tyr Arg Pro
 1               5                   10                  15
Glu Arg Tyr Glu Glu Lys Asp Phe Asp Glu Phe Trp Glu Glu Thr Leu
            20                  25                  30
Ala Glu Ser Glu Lys Phe Pro Leu Asp Pro Val Phe Glu Arg Met Glu
        35                  40                  45
Ser His Leu Lys Thr Val Glu Ala Tyr Asp Val Thr Phe Ser Gly Tyr
    50                  55                  60
Arg Gly Gln Arg Ile Lys Gly Trp Leu Leu Val Pro Lys Leu Glu Glu
65                  70                  75                  80
Glu Lys Leu Pro Cys Val Val Gln Tyr Ile Gly Tyr Asn Gly Gly Arg
                85                  90                  95
Gly Phe Pro His Asp Trp Leu Phe Trp Pro Ser Met Gly Tyr Ile Cys
            100                 105                 110
Phe Val Met Asp Thr Arg Gly Gln Gly Ser Gly Trp Leu Lys Gly Asp
        115                 120                 125
Thr Pro Asp Tyr Pro Glu Gly Pro Val Asp Pro Gln Tyr Pro Gly Phe
    130                 135                 140
Met Thr Arg Gly Ile Leu Asp Pro Arg Thr Tyr Tyr Tyr Arg Arg Val
145                 150                 155                 160
Phe Thr Asp Ala Val Arg Ala Val Glu Ala Ala Ala Ser Phe Pro Gln
                165                 170                 175
Val Asp Gln Glu Arg Ile Val Ile Ala Gly Gly Ser Gln Gly Gly Gly
            180                 185                 190
Ile Ala Leu Ala Val Ser Ala Leu Ser Lys Lys Ala Lys Ala Leu Leu
        195                 200                 205
Cys Asp Val Pro Phe Leu Cys His Phe Arg Arg Ala Val Gln Leu Val
    210                 215                 220
Asp Thr His Pro Tyr Ala Glu Ile Thr Asn Phe Leu Lys Thr His Arg
```

```
225                 230                 235                 240
Asp Lys Glu Glu Ile Val Phe Arg Thr Leu Ser Tyr Phe Asp Gly Val
                245                 250                 255
Asn Phe Ala Ala Arg Ala Lys Ile Pro Ala Leu Phe Ser Val Gly Leu
            260                 265                 270
Met Asp Asn Ile Cys Pro Pro Ser Thr Val Phe Ala Ala Tyr Asn Tyr
        275                 280                 285
Tyr Ala Gly Pro Lys Glu Ile Arg Ile Tyr Pro Tyr Asn Asn His Glu
    290                 295                 300
Gly Gly Gly Ser Phe Gln Ala Val Glu Gln Val Lys Phe Leu Lys Lys
305                 310                 315                 320
Leu Phe Glu Lys Gly
                325
```

(2) INFORMATION FOR SEQ ID NO:62:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 210 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

```
Leu Lys Tyr Phe Lys Ala Arg Leu Ala Gly Ile Thr Leu Leu Gly Leu
1               5                   10                  15
Leu Ala Cys Thr Ser Ala Ser Ala Gln Thr Glu Pro Ile Val Phe Val
            20                  25                  30
His Gly Tyr Ser Gly Ser Ala Ser Asn Trp Asp Thr Met Leu Gly Arg
        35                  40                  45
Phe Arg Ser Asn Gly Tyr Ala Ser Gly Ser Leu Tyr Thr Phe Asn Tyr
    50                  55                  60
Asn Ser Leu Val Ser Ser Asn Arg Thr Ser Ala Ser Glu Leu Arg Ser
65                  70                  75                  80
Phe Val Asn Thr Val Arg Ser Arg His Gly Asn Ala Arg Ile Ala Leu
                85                  90                  95
Val Ala His Ser Asn Gly Gly Leu Val Ser Arg Trp Tyr Arg Ala Glu
            100                 105                 110
Leu Gly Gly Glu Thr Ala Thr Arg Arg Phe Val Thr Leu Gly Thr Pro
        115                 120                 125
His Arg Gly Thr Thr Trp Ala Tyr Ala Cys Tyr Ser Pro Ala Cys Phe
    130                 135                 140
Glu Met Arg Pro Gly Ser Ser Leu Leu Thr Thr Leu Gly Ser Arg Ala
145                 150                 155                 160
Cys Asp Arg Ser Leu Trp Ser Asn Thr Asp Gly Ile Ile Leu Pro Ala
                165                 170                 175
Ser Ser Ala Gln Cys Gly Val Ser Thr Arg Thr Ala Asp Val Ser His
            180                 185                 190
Leu Asp Leu Leu Thr Asp Ser Arg Val Tyr Thr Gln Leu Arg Thr Gln
        195                 200                 205
Leu Gln
    210
```

Claims

1. A polynucleotide comprising a member selected from the group consisting of:

   a) a polynucleotide comprising a nucleic acid sequence set forth in SEQ ID NO:23;
   b) a polynucleotide which encodes an enzyme comprising an amino acid sequence set forth in SEQ ID NO:33;
   c) a polynucleotide comprising at least 15 consecutive bases of a polynucleotide of (a) or (b);
   d) a polynucleotide having at least 70%, 80%, 90%, 95% or 97% identity to the polynucleotide of (a) or (b); and encoding an enzyme having esterase activity; and
   e) a polynucleotide which is complementary to a polynucllеotide of any one of (a) to (d).

2. A polynucleotide hybridizing under stringent conditions to:

   a) a polynucleotide comprising a nucleic acid sequence set forth in SEQ ID NO:23; or
   b) a polynucleotide which is complementary to a polynucleotide of (a);

   and encoding an enzyme having esterase activity.

3. The polynucleotide of claim 1 or 2, wherein the polynucleotide is DNA or RNA.

4. A vector comprising the polynucleotide of any one of claims 1 to 3.

5. A host cell comprising the vector of claim 4.

6. A process for producing a polypeptide comprising expressing from the host cell of claim 5 a polypeptide encoded by said polynucleotide.

7. A process for producing a cell comprising transforming or transfecting the cell with the vector of claim 4 such that the cell expresses the polypeptide encoded by the polynucleotide contained in the vector.

8. An enzyme comprising a member selected from the group consisting of:

   (a) an enzyme encoded by the polynucleotide of any one of claims 1(a), (b) or (d);
   (b) an enzyme comprising an amino acid sequence which is at least 70%, 90% or 95% identical to an amino acid sequence set forth in SEQ ID NO:33 ;
   (c) an enzyme which comprises at least 30 amino acids of the enzyme o (a) or (b); and
   (d) an enzyme obtainable by the process of claim 6.

9. An antibody specifically recognizing the enzyme of claim 8.

10. A method for transferring an amino group from an amino acid to an α-keto acid comprising contacting an amino acid in the presence of an α-keto acid with the enzyme of claim 8.

11. A method for the isolation of the enzyme of claim 8 comprising using the antibody of claim 9.

12. Use of the enzyme of claim 8:

    (a) as ripening starter for cheese in the dairy industry;
    (b) for lignin removal in the pulp and paper industry;
    (c) for the synthesis of carbohydrate derivatives; and/or
    (d) for the conversion of lignocellulosic wastes to fermentable sugars.

13. Use of the polynucleotide of claim 1 or 2 for the identification of similar sequences in other organisms.

Patentansprüche

1. Polynucleotid, umfassend ein Mitglied ausgewählt aus der Gruppe bestehend aus:

(a) einem Polynucleotid, umfassend eine Nucleinsäuresequenz wie in SEQ ID NR:23 dargestellt;
(b) einem Polynucleotid, das ein Enzym codiert, umfassend eine Aminosäuresequenz wie in SEQ ID NR:33 dargestellt;
(c) einem Polynucleotid, umfassend mindestens 15 aufeinanderfolgende Basen eines Polynucleotids aus (a) oder (b);
(d) einem Polynucleotid, das mindestens 70%, 80%, 90%, 95% oder 97% Identität mit dem Polynucleotid aus (a) oder (b) hat; und ein Enzym mit Esterase-Aktivität codiert; und
(e) einem Polynucleotid, das zu einem Polynucleotid aus (a) bis (d) komplementär ist.

2. Polynucleotid, das unter stringenten Bedingungen an:

(a) ein Polynucleotid, umfassend eine Nucleinsäuresequenz wie in SEQ ID NR:23 dargestellt; oder
(b) ein Polynucleotid, das zu einem Polynucleotid aus (a) komplementär ist; hybridisiert und ein Enzym mit Esterase-Aktivität codiert.

3. Polynucleotid nach Anspruch 1 oder 2, wobei das Polynucleotid DNA oder RNA ist.

4. Vektor, umfassend das Polynucleotid nach einem der Ansprüche 1 bis 3.

5. Wirtszelle, umfassend den Vektor nach Anspruch 4.

6. Verfahren zur Herstellung eines Polypeptids, umfassend das Exprimieren eines Polypeptids, das von dem Polynucleotid codiert wird, von der Wirtszelle nach Anspruch 5.

7. Verfahren zur Herstellung einer Zelle, umfassend das Transformieren oder Transfizieren der Zelle mit dem Vektor nach Anspruch 4, so dass die Zelle das Polypeptid exprimiert, das von dem im Vektor enthaltenen Polynucleotid codiert wird.

8. Enzym, umfassend ein Mitglied ausgewählt aus der Gruppe bestehend aus:

(a) einem Enzym, das von dem Polynucleotid nach einem der Ansprüche 1 (a), (b) oder (d) codiert wird;
(b) einem Enzym, umfassend eine Aminosäuresequenz, die zu mindestens 70%, 90% oder 95% identisch ist zu einer Aminosäuresequenz wie in SEQ ID NR:33 dargestellt;
(c) einem Enzym, das mindestens 30 Aminosäuren des Enzyms aus (a) oder (b) umfasst; und
(d) einem Enzym, erhältlich durch das Verfahren nach Anspruch 6.

9. Antikörper, der spezifisch das Enzym nach Anspruch 8 erkennt.

10. Verfahren zum Transferieren einer Aminogruppe von einer Aminosäure zu einer α-Ketosäure, umfassend das Inkontaktbringen einer Aminosäure mit dem Enzym nach Anspruch 8 in Anwesenheit einer α-Ketosäure.

11. Verfahren zum Isolieren des Enzyms nach Anspruch 8, umfassend die Verwendung des Antikörpers nach Anspruch 9.

12. Verwendung des Enzyms nach Anspruch 8:

(a) als Reifestarter für Käse in der Milchindustrie;
(b) zur Ligninentfernung in der Zellstoffindustrie;
(c) zur Synthese von Kohlenhydratderivaten; und/oder
(d) zur Umwandlung von Lignozelluloseabfällen zu fermentierbaren Zuckern.

13. Verwendung des Polynucleotids nach Anspruch 1 oder 2 zur Identifizierung von ähnlichen Sequenzen in anderen Organismen.

Revendications

1. Polynucléotide comprenant un membre choisi dans le groupe constitué de :

a) un polynucléotide comprenant une séquence d'acide nucléique représentée dans SEQ ID NO : 23 ;
b) un polynucléotide qui code une enzyme comprenant une séquence d'acides aminés représentée dans SEQ ID NO : 33 ;
c) un polynucléotide comprenant au moins 15 bases consécutives d'un polynucléotide de a) ou b) ;
d) un polynucléotide ayant au moins 70%, 80%, 90%, 95% ou 97% d'identité avec le polynucléotide de a) ou b) ; et codant une enzyme ayant une activité estérase ; et
e) un polynucleotide complémentaire de l'un quelconque des polynucléotides de a) à d).

2. Polynucléotide qui hybride dans des conditions stringentes à :

a) un polynucléotide comprenant une séquence d'acide nucléique représentée dans SEQ ID NO : 23 ; ou
b) un polynucléotide complémentaire d'un polynucléotide de a) ;

et codant une enzyme ayant une activité estérase.

3. Polynucléotide selon la revendication 1 ou 2, dans lequel le polynucléotide est de l' ADN ou de l'ARN.

4. Vecteur comprenant le polynucléotide selon l'une quelconque des revendications 1 à 3.

5. Cellule hôte comprenant le vecteur selon la revendication 4.

6. Procédé de production d'un polypeptide comprenant l'expression, à partir de la cellule hôte selon la revendication 5, d'un polypeptide codé par ledit polynucléotide.

7. Procédé de production d'une cellule comprenant la transformation ou la transfection de la cellule avec le vecteur selon la revendication 4 de sorte que la cellule exprime le polypeptide codé par le polynucléotide contenu dans le vecteur.

8. Enzyme comprenant un membre choisi dans le groupe constitué de :

a) une enzyme codée par le polynucléotide selon l'une quelconque des revendications 1(a), (b) ou (d) ;
b) une enzyme comprenant une séquence d'acides aminés identique à au moins 70%, 90% ou 95% à une séquence d'acides aminés représentée dans SEQ ID NO : 33 ;
c) une enzyme comprenant au moins 30 acides aminés de l'enzyme de (a) ou (b) ; et
d) une enzyme susceptible d'être obtenue par le procédé de la revendication 6.

9. Anticorps reconnaissant spécifiquement l'enzyme selon la revendication 8.

10. Méthode pour transférer un groupe amino d'un acide aminé à un α-céto acide, comprenant la mise en contact d'un acide aminé en présence d'un α-céto acide avec l'enzyme selon la revendication 8.

11. Méthode pour isoler l'enzyme selon la revendication 8, comprenant l'utilisation de l'anticorps selon la revendication 9.

12. Utilisation de l'enzyme selon la revendication 8 :

a) comme inducteur de la maturation du fromage dans l'industrie laitière ;
b) pour éliminer la lignine dans l'industrie de la pâte et du papier ;
c) pour la synthèse de dérivés d'hydrates de carbone ; et/ou
d) pour la conversion de déchets ligno-cellulosiques en sucres fermentables.

13. Utilisation du polynucléotide de la revendication 1 ou 2 pour l'identification de séquences similaires dans d'autres organismes.

FIGURE 1

***Staphylothermus marinus* - F1-12LC**

```
ATG TCT TTA AAC AAG CAC TCT TGG ATG GAT ATG ATA ATA TTT ATT CTC AGC TTT TCT TTC    60
Met Ser Leu Asn Lys His Ser Trp Met Asp Met Ile Ile Phe Ile Leu Ser Phe Ser Phe    20

CCA TTA ACA ATG ATC GCA TTA GCT ATC TCT ATG TCG TCA TGG TTT AAT ATA TGG AAT AAT   120
Pro Leu Thr Met Ile Ala Leu Ala Ile Ser Met Ser Ser Trp Phe Asn Ile Trp Asn Asn    40

GCA TTA AGC GAT CTA GGA CAT GCT GTT AAA AGC AGT GTT GCT CCA ATA TTC AAT CTA GGT   180
Ala Leu Ser Asp Leu Gly His Ala Val Lys Ser Ser Val Ala Pro Ile Phe Asn Leu Gly    60

CTT GCA ATT GGT GGG ATA CTA ATT GTT ATA GTT GGT TTA AGA AAT CTT TAT TCG TGG AGT   240
Leu Ala Ile Gly Gly Ile Leu Ile Val Ile Val Gly Leu Arg Asn Leu Tyr Ser Trp Ser    80

AGA GTT AAA GGA TCT TTA ATC ATA TCC ATG GGT GTA TTT CTT AAC TTA ATA GGG GTT TTC   300
Arg Val Lys Gly Ser Leu Ile Ile Ser Met Gly Val Phe Leu Asn Leu Ile Gly Val Phe   100

GAC GAA GTA TAT GGT TGG ATA CAT TTC CTA GTC TCA GTA TTG TTT TTC TTA TCA ATA ATA   360
Asp Glu Val Tyr Gly Trp Ile His Phe Leu Val Ser Val Leu Phe Phe Leu Ser Ile Ile   120

GCA TAT TTC ATA GCT ATA TCA ATA CTT GAC AAA TCA TGG ATA GCT GTT CTA CTA ATA ATA   420
Ala Tyr Phe Ile Ala Ile Ser Ile Leu Asp Lys Ser Trp Ile Ala Val Leu Leu Ile Ile   140

GGT CAT ATT GCA ATG TGG TAT CTA CAC TTT GCT TCA GAG ATT CCG AGA GGT GCG GCT ATT   480
Gly His Ile Ala Met Trp Tyr Leu His Phe Ala Ser Glu Ile Pro Arg Gly Ala Ala Ile   160

CCC GAG TTA TTA GCG GTA TTC TCG TTT TTA CCA TTC TAT ATA AGA GAC TAT TTT AAA TCA   540
Pro Glu Leu Leu Ala Val Phe Ser Phe Leu Pro Phe Tyr Ile Arg Asp Tyr Phe Lys Ser   180

TAC ACT AAA CGA TAG                                                               555
Tyr Thr Lys Arg                                                                   185
```

FIGURE 2

Pyrodictium - TAG11-17LC

```
ATG AAA CTC CTT GAG CCC ACA AAT ACC TCC TAC ACG CTG TTA CAG GAT TTA GCA TTG CAT    60
Met Lys Leu Leu Glu Pro Thr Asn Thr Ser Tyr Thr Leu Leu Gln Asp Leu Ala Leu His    20

TTT GCA TTT TAC TGG TTT CTG GCC GTG TAT ACG TGG TTA CCC GGT GTC CTA GTC CGG GGC   120
Phe Ala Phe Tyr Trp Phe Leu Ala Val Tyr Thr Trp Leu Pro Gly Val Leu Val Arg Gly    40

GTA GCT GTG GAC ACA GGG GTG GCT CGG GTG CCT GGG CTC GGC CGG CGC GGT AAG AGG CTG   180
Val Ala Val Asp Thr Gly Val Ala Arg Val Pro Gly Leu Gly Arg Arg Gly Lys Arg Leu    60

CTC CTG GCC GCT GTG GCT GTC TTG GCG CTT GTT GTG TCC GTT GTT GTC CCG GCT TAT GTG   240
Leu Leu Ala Ala Val Ala Val Leu Ala Leu Val Val Ser Val Val Val Pro Ala Tyr Val    80

GCG TAT AGT AGT CTG CAC CCG GAG AGC TGT CGG CCC GTT GCG CCG GAG GGG CTC ACC TAC   300
Ala Tyr Ser Ser Leu His Pro Glu Ser Cys Arg Pro Val Ala Pro Glu Gly Leu Thr Tyr   100

AAA GAG TTC AGC GTG ACC GCG GAG GAT GGC TTG GTG GTT CGG GGC TGG GTG CTG GGC CCC   360
Lys Glu Phe Ser Val Thr Ala Glu Asp Gly Leu Val Val Arg Gly Trp Val Leu Gly Pro   120

GC GCT GGG GGC AAC CCG GTG TTC GTT TTG ATG CAC GGG TAT ACT GGG TGC CGC TCG GCG    420
Gly Ala Gly Gly Asn Pro Val Phe Val Leu Met His Gly Tyr Thr Gly Cys Arg Ser Ala   140

CCC TAC ATG GCT GTG CTG GCC CGG GAG CTC GTG GAG TGG GGG TAC CCG GTG GTT GTG TTC   480
Pro Tyr Met Ala Val Leu Ala Arg Glu Leu Val Glu Trp Gly Tyr Pro Val Val Val Phe   160

GAC TTC CGG GGC CAC GGG GAG AGC GGG GGC TCG ACG ACG ATT GGG CCC CGG GAG GTG CTG   540
Asp Phe Arg Gly His Gly Glu Ser Gly Gly Ser Thr Thr Ile Gly Pro Arg Glu Val Leu   180

GAT GCC CGG GCT GTG GTG GGC TAT GTC TCG GAG CGG TTC CCC GGC CGC CGG ATA ATA TTG   600
Asp Ala Arg Ala Val Val Gly Tyr Val Ser Glu Arg Phe Pro Gly Arg Arg Ile Ile Leu   200

GTG GGG TTC AGT ATG GGC GGC GCT GTA GCG ATC GTG GAG GGT GCT GGG GAC CCG CGG GTC   660
Val Gly Phe Ser Met Gly Gly Ala Val Ala Ile Val Glu Gly Ala Gly Asp Pro Arg Val   220

TAC GCG GTG GCT GCT GAT AGC CCG TAC TAT AGG CTC CGG GAC GTC ATA CCC CGG TGG CTG   720
Tyr Ala Val Ala Ala Asp Ser Pro Tyr Tyr Arg Leu Arg Asp Val Ile Pro Arg Trp Leu   240

GAG TAC AAG ACG CCG CTG CCG GGC TGG GTG GGT GTG CTG GCC GGG TTC TAC GGG AGG CTG   780
Glu Tyr Lys Thr Pro Leu Pro Gly Trp Val Gly Val Leu Ala Gly Phe Tyr Gly Arg Leu   260

ATG GCG GGC GTT GAC CTC GGC TTC GGC CCC GCT GGG GTG GAG CGC GTG GAT AAG CCG TTG   840
Met Ala Gly Val Asp Leu Gly Phe Gly Pro Ala Gly Val Glu Arg Val Asp Lys Pro Leu   280

CTG GTG GTG TAT GGG CCC CGG GAC CCG CTG GTG ACG CGG GAC GAG GCG AGG AGC CTG GCG   900
Leu Val Val Tyr Gly Pro Arg Asp Pro Leu Val Thr Arg Asp Glu Ala Arg Ser Leu Ala   300

TCC CGT AGC CCG TGT GGC CGT CTC GTC GAG GTT CCT GGG GCT GGC CAC GTG GAG GCC GTG   960
Ser Arg Ser Pro Cys Gly Arg Leu Val Glu Val Pro Gly Ala Gly His Val Glu Ala Val   320

GAT GTG CTC GGG CCG GGC CGC TAC GCA GAC ATG CTG ATA GAG CTG GCG CAC GAG GAG TGC  1020
Asp Val Leu Gly Pro Gly Arg Tyr Ala Asp Met Leu Ile Glu Leu Ala His Glu Glu Cys   340

CCT CCG GGG GCC GGT GGC TGA                                                      1041
Pro Pro Gly Ala Gly Gly                                                           346
```

FIGURE 3

***Archaeoglobus Venificus* SN P6-24LC**

```
ATG CCA TAT GTT AGG AAT GGT GGT GTA AAT ATC TAT TAT GAA CTG GTG GAT GGA CCT GAG   60
Met Pro Tyr Val Arg Asn Gly Gly Val Asn Ile Tyr Tyr Glu Leu Val Asp Gly Pro Glu   20

CCA CCA ATT GTC TTT GTT CAC GGA TGG ACA GCA AAT ATG AAT TTT TGG AAA GAG CAA AGA  120
Pro Pro Ile Val Phe Val His Gly Trp Thr Ala Asn Met Asn Phe Trp Lys Glu Gln Arg   40

CGT TAT TTT GCA GGC AGG AAT ATG ATG TTG TTT GTC GAT AAC AGA GGT CAT GGC AGG TCC  180
Arg Tyr Phe Ala Gly Arg Asn Met Met Leu Phe Val Asp Asn Arg Gly His Gly Arg Ser   60

GAT AAG CCA CTT GGA TAC GAT TTC TAC AGA TTT GAG AAC TTC ATT TCA GAT TTA GAT GCG  240
Asp Lys Pro Leu Gly Tyr Asp Phe Tyr Arg Phe Glu Asn Phe Ile Ser Asp Leu Asp Ala   80

GTT GTT AGG GAG ACT GGA GTG GAG AAA TTT GTT CTC GTC GGA CAT TCA TTC GGA ACA ATG  300
Val Val Arg Glu Thr Gly Val Glu Lys Phe Val Leu Val Gly His Ser Phe Gly Thr Met  100

ATC TCT ATG AAG TAC TGT TCG GAG TAT CGG AAT CGG GTT CTT GCT CTA ATC CTC ATA GGT  360
Ile Ser Met Lys Tyr Cys Ser Glu Tyr Arg Asn Arg Val Leu Ala Leu Ile Leu Ile Gly  120

GGT GGG AGC AGA ATA AAG CTT CTA CAC AGA ATT GGA TAT CCT TTA GCA AAG ATT CTT GCA  420
Gly Gly Ser Arg Ile Lys Leu Leu His Arg Ile Gly Tyr Pro Leu Ala Lys Ile Leu Ala  140

TCC ATT GCA TAC AAG AAG TCT TCA AGA TTG GTC GCA GAT CTT TCC TTT GGC AAA AAT GCT  480
Ser Ile Ala Tyr Lys Lys Ser Ser Arg Leu Val Ala Asp Leu Ser Phe Gly Lys Asn Ala  160

GGT GAA CTT AAA GAG TGG GGA TGG AAA CAG GCA ATG GAT TAT ACA CCC TCC TAC GTG GCA  540
Gly Glu Leu Lys Glu Trp Gly Trp Lys Gln Ala Met Asp Tyr Thr Pro Ser Tyr Val Ala  180

ATG TAC ACG TAC AGA ACT CTA ACG AAA GTG AAT CTT GAA AAT ATC TTG GAG AAA ATA GAC  600
Met Tyr Thr Tyr Arg Thr Leu Thr Lys Val Asn Leu Glu Asn Ile Leu Glu Lys Ile Asp  200

TGT CCA ACA CTG ATT ATC GTT GGA GAA GAG GAT GCA CTA TTG CCC GTT AGC AAA TCA GTT  660
Cys Pro Thr Leu Ile Ile Val Gly Glu Glu Asp Ala Leu Leu Pro Val Ser Lys Ser Val  220

GAG CTG AGC AGG AGG ATA GAA AAC TCA AAG CTT GTG ATC ATC CCA AAC TCG GGG CAT TGC  720
Glu Leu Ser Arg Arg Ile Glu Asn Ser Lys Leu Val Ile Ile Pro Asn Ser Gly His Cys  240

GTA ATG CTT GAG AGT CCA AGT GAG GTT AAT AGA GCA ATG GAC GAA TTC ATT TCT TCA GCA  780
Val Met Leu Glu Ser Pro Ser Glu Val Asn Arg Ala Met Asp Glu Phe Ile Ser Ser Ala  260

CAG TTC TAA                                                                      789
Gln Phe                                                                          263
```

FIGURE 4

Aquifex pyrophilus - 28LC

```
TTG AGA TTG AGG AAA TTT GAA GAG ATA AAC CTC GTT CTT TCG GGA GGA GCT GCA AAG GGC    60
Leu Arg Leu Arg Lys Phe Glu Glu Ile Asn Leu Val Leu Ser Gly Gly Ala Ala Lys Gly    20

ATA GCC CAC ATA GGT GTT TTG AAA GCT ATA AAC GAG CTC GGT ATA AGG GTG AGG GCT TTA   120
Ile Ala His Ile Gly Val Leu Lys Ala Ile Asn Glu Leu Gly Ile Arg Val Arg Ala Leu    40

AGC GGG GTG AGC GCC GGG GCA ATC GTT TCG GTC TTT TAT GCC TCA GGC TAC TCC CCT GAA   180
Ser Gly Val Ser Ala Gly Ala Ile Val Ser Val Phe Tyr Ala Ser Gly Tyr Ser Pro Glu    60

GGG ATG TTC AGC CTT CTG AAG AGG GTA AAC TGG CTG AAG CTG TTT AAG TTC AAG CCA CCT   240
Gly Met Phe Ser Leu Leu Lys Arg Val Asn Trp Leu Lys Leu Phe Lys Phe Lys Pro Pro    80

CTG AAG GGA TTG ATA GGG TGG GAG AAG GCT ATA AGA TTC CTT GAG GAA GTT CTC CCT TAC   300
Leu Lys Gly Leu Ile Gly Trp Glu Lys Ala Ile Arg Phe Leu Glu Glu Val Leu Pro Tyr   100

AGG AGA ATA GAA AAA CTT GAG ATA CCG ACG TAT ATA TGC GCG ACG GAT TTA TAC TCG GGA   360
Arg Arg Ile Glu Lys Leu Glu Ile Pro Thr Tyr Ile Cys Ala Thr Asp Leu Tyr Ser Gly   120

AGG GCT CTA TAC CTC TCG GAA GGG AGT TTA ATC CCC GCA CTT CTC GGC AGC TGT GCA ATT   420
Arg Ala Leu Tyr Leu Ser Glu Gly Ser Leu Ile Pro Ala Leu Leu Gly Ser Cys Ala Ile   140

CCC GGC ATA TTT GAA CCC GTT GAG TAT AAG AAT TAC TTG CTC GTT GAC GGA GGT ATA GTT   480
Pro Gly Ile Phe Glu Pro Val Glu Tyr Lys Asn Tyr Leu Leu Val Asp Gly Gly Ile Val   160

AAC AAC CTT CCC GTT GAG CCC TTT CAG GAA AGC GGT ATT CCC ACC GTT TGC GTT GAT GTC   540
Asn Asn Leu Pro Val Glu Pro Phe Gln Glu Ser Gly Ile Pro Thr Val Cys Val Asp Val   180

CTT CCC ATA GAG CCG GAA AAG GAT ATA AAG AAC ATT CTT CAC ATC CTT TTG AGG AGC TTC   600
Leu Pro Ile Glu Pro Glu Lys Asp Ile Lys Asn Ile Leu His Ile Leu Leu Arg Ser Phe   200

TTT CTT GCG GTC CGC TCA AAC TCC GAA AAG AGA AAG GAG TTT TGT GAC CTC GTT ATA GTT   660
Phe Leu Ala Val Arg Ser Asn Ser Glu Lys Arg Lys Glu Phe Cys Asp Leu Val Ile Val   220

CCT GAG CTT GAG GAG TTC ACA CCC CTT GAT GTT AGA AAA GCG GAC CAA ATA ATG GAG AGG   720
Pro Glu Leu Glu Glu Phe Thr Pro Leu Asp Val Arg Lys Ala Asp Gln Ile Met Glu Arg   240

GGA TAC ATA AAG GCC TTA GAG GTA CTT TCT GAA TAG                                   756
Gly Tyr Ile Lys Ala Leu Glu Val Leu Ser Glu                                       252
```

FIGURE 5

M11TL-29L.

```
ATG TTT AAT ATC AAT GTC TTT GTT AAT ATA TCT TGG CTG TAT TTT TCA GGG ATA GTT ATG    60
Met Phe Asn Ile Asn Val Phe Val Asn Ile Ser Trp Leu Tyr Phe Ser Gly Ile Val Met    20

AAG ACT GTG GAA GAG TAT GCG CTA CTT GAA ACA GGC GTA AGA GTG TTT TAT CGG TGT GTA   120
Lys Thr Val Glu Glu Tyr Ala Leu Leu Glu Thr Gly Val Arg Val Phe Tyr Arg Cys Val    40

ATC CCG GAG AAA GCT TTT AAC ACT TTG ATA ATA GGT TCA CAC GGA TTG GGG GCG CAC AGT   180
Ile Pro Glu Lys Ala Phe Asn Thr Leu Ile Ile Gly Ser His Gly Leu Gly Ala His Ser    60

GGA ATC TAC ATT AGT GTT GCT GAA GAA TTT GCT AGG CAC GGA TTT GGA TTC TGC ATG CAC   240
Gly Ile Tyr Ile Ser Val Ala Glu Glu Phe Ala Arg His Gly Phe Gly Phe Cys Met His    80

GAT CAA AGG GGA CAT GGG AGA ACG GCA AGC GAT AGA GAA AGA GGG TAT GTG GAG GGC TTT   300
Asp Gln Arg Gly His Gly Arg Thr Ala Ser Asp Arg Glu Arg Gly Tyr Val Glu Gly Phe   100

CAC AAC TTC ATA GAG GAT ATG AAG GCC TTC TCC GAT TAT GCC AAG TGG CGC GTG GGA GGT   360
His Asn Phe Ile Glu Asp Met Lys Ala Phe Ser Asp Tyr Ala Lys Trp Arg Val Gly Gly   120

GAC GAA ATA ATA TTG CTA GGA CAC AGT ATG GGC GGG CTG ATA GCG CTC TTA ACA GTT GCA   420
Asp Glu Ile Ile Leu Leu Gly His Ser Met Gly Gly Leu Ile Ala Leu Leu Thr Val Ala   140

ACT TAT AAA GAA ATC GCC AAG GGA GTT ATC GCG CTA GCC CCG GCC CTC CAA ATC CCC TTA   480
Thr Tyr Lys Glu Ile Ala Lys Gly Val Ile Ala Leu Ala Pro Ala Leu Gln Ile Pro Leu   160

ACC CCG GCT AGA AGA CTT GTT CTA AGC CTC GCG TCA AGG CTT GCC CCG CAT TCT AAG ATC   540
Thr Pro Ala Arg Arg Leu Val Leu Ser Leu Ala Ser Arg Leu Ala Pro His Ser Lys Ile   180

ACC TTA CAA AGG AGA TTG CCG CAG AAA CCA GAG GGT TTT CAA AGA GCA AAA GAT ATA GAA   600
Thr Leu Gln Arg Arg Leu Pro Gln Lys Pro Glu Gly Phe Gln Arg Ala Lys Asp Ile Glu   200

TAC AGT CTG AGT GAA ATA TCA GTC AAG CTC GTG GAC GAA ATG ATT AAA GCA TCA TCT ATG   660
Tyr Ser Leu Ser Glu Ile Ser Val Lys Leu Val Asp Glu Met Ile Lys Ala Ser Ser Met   220

TTC TGG ACC ATA GCA GGG GAA ATT AAT ACT CCC GTC CTG CTT ATT CAT GGG GAA AAA GAC   720
Phe Trp Thr Ile Ala Gly Glu Ile Asn Thr Pro Val Leu Leu Ile His Gly Glu Lys Asp   240

AAT GTC ATA CCT CCG GAG GCG AGC AAA AAA GCC TAC CAA TTA ATA CCT TCA TTC CCT AAA   780
Asn Val Ile Pro Pro Glu Ala Ser Lys Lys Ala Tyr Gln Leu Ile Pro Ser Phe Pro Lys   260

GAG TTG AAA ATA TAC CCC GAT CTT GGA CAC AAC TTG TTT TTT GAA CCA GGC GCG GTG AAA   840
Glu Leu Lys Ile Tyr Pro Asp Leu Gly His Asn Leu Phe Phe Glu Pro Gly Ala Val Lys   280

ATC GTC ACA GAC ATT GTA GAG TGG GTT AAG AAT CTA CCC AGG GAA AAT CCT TAA           894
Ile Val Thr Asp Ile Val Glu Trp Val Lys Asn Leu Pro Arg Glu Asn Pro               298
```

FIGURE 6

Thermococcus CL-2-30LC

```
ATG GAG GTT TAC AAG GCC AAA TTC GGC GAA GCA AAG CTC GGC TGG GTC GTT CTG GTT CAT    60
Met Glu Val Tyr Lys Ala Lys Phe Gly Glu Ala Lys Leu Gly Trp Val Val Leu Val His    20

GGC CTC GGC GAG CAC AGC GGA AGG TAT GGA AGA CTG ATT AAG GAA CTC AAC TAT GCC GGC   120
Gly Leu Gly Glu His Ser Gly Arg Tyr Gly Arg Leu Ile Lys Glu Leu Asn Tyr Ala Gly    40

TTT GGA GTT TAC ACC TTC GAC TGG CCC GGC CAC GGG AAG AGC CCG GGC AAG AGA GGG CAC   180
Phe Gly Val Tyr Thr Phe Asp Trp Pro Gly His Gly Lys Ser Pro Gly Lys Arg Gly His    60

ACG AGC GTC GAG GAG GCG ATG GAA ATC ATC GAC TCG ATA ATC GAG GAG ATC AGG GAG AAG   240
Thr Ser Val Glu Glu Ala Met Glu Ile Ile Asp Ser Ile Ile Glu Glu Ile Arg Glu Lys    80

CCC TTC CTC TTC GGC CAC AGC CTC GGT GGT CTA ACT GTC ATC AGG TAC GCT GAG ACG CGG   300
Pro Phe Leu Phe Gly His Ser Leu Gly Gly Leu Thr Val Ile Arg Tyr Ala Glu Thr Arg   100

CCC GAT AAA ATA CGG GGA TTA ATA GCT TCC TCG CCT GCC CTC GCC AAG AGC CCG GAA ACG   360
Pro Asp Lys Ile Arg Gly Leu Ile Ala Ser Ser Pro Ala Leu Ala Lys Ser Pro Glu Thr   120

CCG GGC TTC ATG GTG GCC CTC GCG AAG TTC CTT GGA AAG ATC GCC CCG GGA GTT GTT CTC   420
Pro Gly Phe Met Val Ala Leu Ala Lys Phe Leu Gly Lys Ile Ala Pro Gly Val Val Leu   140

TCC AAC GGC ATA AAG CCG GAA CTC CTC TCG AGG AAC AGG GAC GCC GTG AGG AGG TAC GTT   480
Ser Asn Gly Ile Lys Pro Glu Leu Leu Ser Arg Asn Arg Asp Ala Val Arg Arg Tyr Val   160

GAA GAC CCA CTC GTC CAC GAC AGG ATT TCG GCC AAG CTG GGA AGG AGC ATC TTC GTG AAC   540
Glu Asp Pro Leu Val His Asp Arg Ile Ser Ala Lys Leu Gly Arg Ser Ile Phe Val Asn   180

ATG GAG CTG GCC CAC AGG GAG GCG GAC AAG ATA AAA GTC CCG ATC CTC CTT CTG ATC GGC   600
Met Glu Leu Ala His Arg Glu Ala Asp Lys Ile Lys Val Pro Ile Leu Leu Leu Ile Gly   200

ACT GGC GAT GTA ATA ACC CCG CCT GAA GGC TCA CGC AGA CTC TTC GAG GAG CTG GCC GTC   660
Thr Gly Asp Val Ile Thr Pro Pro Glu Gly Ser Arg Arg Leu Phe Glu Glu Leu Ala Val   220

GAG AAC AAA ACC CTG AGG GAG TTC GAG GGG GCG TAC CAC GAG ATA TTT GAA GAC CCC GAG   720
Glu Asn Lys Thr Leu Arg Glu Phe Glu Gly Ala Tyr His Glu Ile Phe Glu Asp Pro Glu   240

TGG GCC GAG GAG TTC CAC GAA ACA ATT GTT AAG TGG CTG GTT GAA AAA TCG TAC TCT TCG   780
Trp Ala Glu Glu Phe His Glu Thr Ile Val Lys Trp Leu Val Glu Lys Ser Tyr Ser Ser   260

GCT CAA TAA                                                                       789
Ala Gln                                                                           263
```

FIGURE 7

Aquifex VF5-34LC

```
TTG ATT GGC AAT TTG AAA TTG AAG AGG TTT GAA GAG GTT AAC TTA GTT CTT TCG GGA GGG    60
Leu Ile Gly Asn Leu Lys Leu Lys Arg Phe Glu Glu Val Asn Leu Val Leu Ser Gly Gly    20

GCT GCC AAG GGT ATC GCC CAT ATA GGT GTT TTA AAA GCT CTG GAA GAG CTC GGT ATA AAG   120
Ala Ala Lys Gly Ile Ala His Ile Gly Val Leu Lys Ala Leu Glu Glu Leu Gly Ile Lys    40

GTA AAG AGG CTC AGC GGG GTA AGT GCT GGA GCT ATC GTT TCC GTC TTT TAC GCT TCG GGC   180
Val Lys Arg Leu Ser Gly Val Ser Ala Gly Ala Ile Val Ser Val Phe Tyr Ala Ser Gly    60

TAC ACT CCC GAC GAG ATG TTA AAA CTC CTG AAA GAG GTA AAC TGG CTC AAA CTT TTT AAG   240
Tyr Thr Pro Asp Glu Met Leu Lys Leu Leu Lys Glu Val Asn Trp Leu Lys Leu Phe Lys    80

TTC AAA ACA CCG AAA ATG GGC TTA ATG GGG TGG GAG AAG GCT GCA GAG TTT TTG GAA AAA   300
Phe Lys Thr Pro Lys Met Gly Leu Met Gly Trp Glu Lys Ala Ala Glu Phe Leu Glu Lys   100

GAG CTC GGA GTT AAG AGG CTG GAA GAC CTG AAC ATA CCA ACC TAT CTT TGC TCG GCG GAT   360
Glu Leu Gly Val Lys Arg Leu Glu Asp Leu Asn Ile Pro Thr Tyr Leu Cys Ser Ala Asp   120

CTG TAC ACG GGA AAG GCT CTT TAC TTC GGC AGA GGT GAC TTA ATT CCC GTG CTT CTC GGA   420
Leu Tyr Thr Gly Lys Ala Leu Tyr Phe Gly Arg Gly Asp Leu Ile Pro Val Leu Leu Gly   140

AGT TGT TCC ATA CCC GGG ATT TTT GAA CCA GTT GAG TAC GAG AAT TTT CTA CTT GTT GAC   480
Ser Cys Ser Ile Pro Gly Ile Phe Glu Pro Val Glu Tyr Glu Asn Phe Leu Leu Val Asp   160

GGA GGT ATA GTG AAC AAC CTG CCC GTA GAA CCT TTG GAA AAG TTC AAA GAA CCC ATA ATC   540
Gly Gly Ile Val Asn Asn Leu Pro Val Glu Pro Leu Glu Lys Phe Lys Glu Pro Ile Ile   180

GGG GTA GAT GTG CTT CCC ATA ACT CAA GAA AGA AAG ATT AAA AAT ATA CTC CAC ATC CTT   600
Gly Val Asp Val Leu Pro Ile Thr Gln Glu Arg Lys Ile Lys Asn Ile Leu His Ile Leu   200

ATA AGG AGC TTC TTT CTG GCG GTT CGT TCC AAT TCG GAA AAG AGA AAG GAG TTC TGC AAC   660
Ile Arg Ser Phe Phe Leu Ala Val Arg Ser Asn Ser Glu Lys Arg Lys Glu Phe Cys Asn   220

GTA GTT ATA GAA CCT CCC CTT GAA GAG TTC TCT CCT CTG GAC GTA AAT AAG GCG GAC GAG   720
Val Val Ile Glu Pro Pro Leu Glu Glu Phe Ser Pro Leu Asp Val Asn Lys Ala Asp Glu   240

ATA TTC TGC GGG GAT ATG AGA GCA CTT TAA                                           750
Ile Phe Cys Gly Asp Met Arg Ala Leu                                               250
```

FIGURE 8

Teredinibacter - 42L

```
ATG CCA GCT AAT GAC TCA CCC ACG ATC GAC TTT AAT CCT CGC GGC ATT CTT CGC AAC GCT    60
Met Pro Ala Asn Asp Ser Pro Thr Ile Asp Phe Asn Pro Arg Gly Ile Leu Arg Asn Ala    20

CAC GCA CAG GTT ATT TTA GCG ACT TCC GGC TTG CGC AAA GCG TTT TTG AAA CGC ACG CAC   120
His Ala Gln Val Ile Leu Ala Thr Ser Gly Leu Arg Lys Ala Phe Leu Lys Arg Thr His    40

AAG AGC TAC CTC AGC ACT GCC CAA TGG CTG GAG CTC GAT GCC GGC AAC GGA GTT ACC TTG   180
Lys Ser Tyr Leu Ser Thr Ala Gln Trp Leu Glu Leu Asp Ala Gly Asn Gly Val Thr Leu    60

GCC GGA GAG CTT AAC ACA GCG CCT GCA ACT GCA TCC TCC TCC CAC CCG GCG CAC AAG AAC   240
Ala Gly Glu Leu Asn Thr Ala Pro Ala Thr Ala Ser Ser Ser His Pro Ala His Lys Asn    80

ACT CTG GTT ATT GTG CTG CAC GGC TGG GAA GGC TCC AGC CAG TCG GCC TAT GCG ACC TCC   300
Thr Leu Val Ile Val Leu His Gly Trp Glu Gly Ser Ser Gln Ser Ala Tyr Ala Thr Ser   100

GCT GGC AGC ACG CTT TTC GAC AAT GGG TTC GAC ACT TTT CGC CTT AAT TTT CGC GAT CAC   360
Ala Gly Ser Thr Leu Phe Asp Asn Gly Phe Asp Thr Phe Arg Leu Asn Phe Arg Asp His   120

GGC GAC ACC TAC CAC TTA AAC CGC GGC ATA TTT AAC TCA TCG CTG ATT GAC GAA GTA GTG   420
Gly Asp Thr Tyr His Leu Asn Arg Gly Ile Phe Asn Ser Ser Leu Ile Asp Glu Val Val   140

GGC GCA GTC AAA GCC ATC CAG CAG CAA ACC GAC TAC GAC AAG TAT TGC CTG ATG GGG TTC   480
Gly Ala Val Lys Ala Ile Gln Gln Gln Thr Asp Tyr Asp Lys Tyr Cys Leu Met Gly Phe   160

TCA CTG GGT GGG AAC TTT GCC TTG CGC GTC GCG GTG CGG GAA CAG CAT CTC GCT AAA CCG   540
Ser Leu Gly Gly Asn Phe Ala Leu Arg Val Ala Val Arg Glu Gln His Leu Ala Lys Pro   180

CTA GCG GGC GTG CTC GCC GTA TGC CCG GTA CTC GAC CCC GCA CAC ACC ATG ATG GCC CTA   600
Leu Ala Gly Val Leu Ala Val Cys Pro Val Leu Asp Pro Ala His Thr Met Met Ala Leu   200

AAC CGA GGT GCG TTT TTC TAC GGC CGC TAT TTT GCG CAT AAA TGG AAG CGC TCG TTA ACC   660
Asn Arg Gly Ala Phe Phe Tyr Gly Arg Tyr Phe Ala His Lys Trp Lys Arg Ser Leu Thr   220

GCA AAA CTT GCA GCT TTC CCA GAC TAC AAA TAC GGC AAA GAT TTA AAA TCG ATA CAC ACG   720
Ala Lys Leu Ala Ala Phe Pro Asp Tyr Lys Tyr Gly Lys Asp Leu Lys Ser Ile His Thr   240

CTT GAT GAG TTA AAC AAC TAT TTC ATT CCC CGC TAC ACC GGC TTC AAC TCA GTC TCC GAA   780
Leu Asp Glu Leu Asn Asn Tyr Phe Ile Pro Arg Tyr Thr Gly Phe Asn Ser Val Ser Glu   260

TAC TTC AAA AGT TAC ACG CTC ACC GGG CAG AAG CTC GCG TTT CTC AAC TGC CCC AGT TAC   840
Tyr Phe Lys Ser Tyr Thr Leu Thr Gly Gln Lys Leu Ala Phe Leu Asn Cys Pro Ser Tyr   280

ATT CTG GCA GCT GGC GAC GAC CCA ATA ATT CCA GCA TCC GAC TTT CAG AAA ATA GCC AAG   900
Ile Leu Ala Ala Gly Asp Asp Pro Ile Ile Pro Ala Ser Asp Phe Gln Lys Ile Ala Lys   300

CCT GCG AAT CTG CAC ATA ACA GTA ACG CAA CAA GGT TCT CAT TGC GCA TAC CTG GAA AAC   960
Pro Ala Asn Leu His Ile Thr Val Thr Gln Gln Gly Ser His Cys Ala Tyr Leu Glu Asn   320

CTG CAT AAA CCT AGT GCT GCC GAC AAA TAT GCG GTG AAA TTA TTT GGA GCC TGT TGA      1017
Leu His Lys Pro Ser Ala Ala Asp Lys Tyr Ala Val Lys Leu Phe Gly Ala Cys End       339
```

FIGURE 9

Archeoglobus fulgidas VC16 - 16MC1

```
ATG CTT GAT ATG CCA ATC GAC CCT GTT TAC TAC CAG CTT GCT GAG TAT
Met Leu Asp Met Pro Ile Asp Pro Val Tyr Tyr Gln Leu Ala Glu Tyr
 1               5                   10                  15

TTC GAC AGT CTG CCG AAG TTC GAC CAG TTT TCC TCG GCC AGA GAG TAC
Phe Asp Ser Leu Pro Lys Phe Asp Gln Phe Ser Ser Ala Arg Glu Tyr
             20                  25                  30

AGG GAG GCG ATA AAT CGA ATA TAC GAG GAG AGA AAC CGG CAG CTG AGC
Arg Glu Ala Ile Asn Arg Ile Tyr Glu Glu Arg Asn Arg Gln Leu Ser
         35                  40                  45

CAG CAT GAG AGG GTT GAA AGA GTT GAG GAC AGG ACG ATT AAG GGG AGG
Gln His Glu Arg Val Glu Arg Val Glu Asp Arg Thr Ile Lys Gly Arg
     50                  55                  60

AAC GGA GAC ATC AGA GTC AGA GTT TAC CAG CAG AAG CCC GAT TCC CCG
Asn Gly Asp Ile Arg Val Arg Val Tyr Gln Gln Lys Pro Asp Ser Pro
 65                  70                  75                  80

GGT CTG GTT TAC TAT CAC GGT GGT GGA TTT GTG ATT TGC AGC ATC GAG
Val Leu Val Tyr Tyr His Gly Gly Gly Phe Val Ile Cys Ser Ile Glu
                 85                  90                  95

TCG CAC GAC GCC TTA TGC AGG AGA AYY GCG AGA CTT TCA AAC TCT ACC
Ser His Asp Ala Leu Cys Arg Arg Ile Ala Arg Leu Ser Asn Ser Thr
            100                 105                 110

GTA GTC TCC GTG GAT TAC AGG CTC GCT CCT GAG CAC AAG TTT CCC CCC
Val Val Ser Val Asp Tyr Arg Leu Ala Pro Glu His Lys Phe Pro Ala
        115                 120                 125

CCA GTT TAT CAT TGC TAC GAT GCG ACC AAG TGG GTT GCT GAG AAC CGG
Ala Val Tyr Asp Cys Tyr Asp Ala Thr Lys Trp Val Ala Glu Asn Ala
    130                 135                 140

GAG GAG CTG AGG ATT GAC CCG TCA AAA ATC TTC GTT GGG GGG GAC AGT
Glu Glu Leu Arg Ile Asp Pro Ser Lys Ile Phe Val Gly Gly Asp Ser
145                 150                 155                 160

GCG GGA CGG AAT CTT GCC CCG GCG CTT TCA ATA ATG GCG AGA GAC AGC
Ala Gly Gly Asn Leu Ala Ala Ala Val Ser Ile Met Ala Arg Asp Ser
                165                 170                 175

GGA GAA GAT TTC ATA AAG CAT CAA ATT CTA ACT TAC CCC GTT GTG AAC
Gly Glu Asp Phe Ile Lys His Gln Ile Leu Ile Tyr Pro Val Val Asn
            180                 185                 190

TTT GTA GCC CCC ACA CCA TCG CTT CTG GAG TTT GGA GAG GGG CTG TGG
Phe Val Ala Pro Thr Pro Ser Leu Leu Glu Phe Gly Glu Gly Leu Trp
        195                 200                 205

ATT CTC GAC CAG AAG ATA ATG AGT TGG TTC TCG GAG CAG TAC TTC TCC
Ile Leu Asp Gln Lys Ile Met Ser Trp Phe Ser Glu Gln Tyr Phe Ser
    210                 215                 230

AGA GAG GAA GAT AAG TTC AAG CCC CTC GCC TCC GTA ATC TTT GCG GAC
Arg Glu Glu Aso Lys Phe Asn Pro Leu Ala Ser Val Ile Phe Ala Asp
235                 240                 245                 250
```

```
CTT GAG AAC CTA CCT CCT GCG CTG ATC ATA ACC GCC GAA TAC GAC CCG
Leu Glu Asn Leu Pro Pro Ala Leu Ile Ile Thr Ala Glu Tyr Asp Pro
                255                 260                 265

CTG AGA GAT GAA GGA GAA GTT TTC GGG CAG ATG CTG AGA AGA GCC GGT
Leu Arg Asp Glu Gly Glu Val Phe Gly Gln Met Leu Arg Arg Ala Gly
            270                 275                 280

GTT GAG GCG AGC ATC GTC AGA TAC AGA GGC GTG CTT CAC GGA TTC ATC
Val Glu Ala Ser Ile Val Arg Tyr Arg Gly Val Leu His Gly Phe Ile
        285                 290                 295

AAT TAC TAT CCC GTG CTG AAG GCT GCG AGG GAT GCG ATA AAC CAG ATT
Asn Tyr Tyr Pro Val Leu Lys Ala Ala Arg Asp Ala Ile Asn Gln Ile
    300                 305                 310

GCC GCT CTT CTT GTG TTC GAC TAG
Ala Ala Leu Leu Val Phe Asp
315                 320
```

FIGURE 10

Sulfolobus Solfataricus P1 - SLC1

```
ATG CCC CTA GAT CCT AGA ATT AAA AAG TTA CTA GAA TCA GCT CTT ACT
Met Pro Leu Asp Pro Arg Ile Lys Lys Leu Leu Glu Ser Ala Leu Thr
1               5                   10                  15

ATA CCA ATT GGT AAA GCC CCA GTA GAA GAG GTA AGA AAG ATA TTT AGG
Ile Pro Ile Gly Lys Ala Pro Val Glu Glu Val Arg Lys Ile Phe Arg
            20                  25                  30

CAA TTA GCG TCG GCA GCT CCC AAA GTC GAA GTT GGA AAA GTA GAA GAT
Gln Leu Ala Ser Ala Ala Pro Lys Val Glu Val Gly Lys Val Glu Asp
        35                  40                  45

ATA AAA ATA CCA GGC AGT GAA ACC GTT ATA AAC GCT AGA GTG TAT TTT
Ile Lys Ile Pro Gly Ser Glu Thr Val Ile Asn Ala Arg Val Tyr Phe
    50                  55                  60

CCG AAG AGT AGC GGT CCT TAT GGT GTT CTA GTG TAT CTT CAT GGA GGC
Pro Lys Ser Ser Gly Pro Tyr Gly Val Leu Val Tyr Leu His Gly Gly
65                  70                  75                  80

GGT TTT GTA ATA GGC GAT GTG GAA TCT TAT GAC CCA TTA TGT AGA GCA
Gly Phe Val Ile Gly Asp Val Glu Ser Tyr Asp Pro Leu Cys Arg Ala
                85                  90                  95

ATT ACA AAT GCG TGC AAT TGC GTT GTA GTA TCA GTG GAC TAT AGG TTA
Ile Thr Asn Ala Cys Asn Cys Val Val Val Ser Val Asp Tyr Arg Leu
            100                 105                 110

GCT CCA GAA TAC AAG TTT CCT TCT GCA GTT ATC GAT TCA TTT GAC GCT
Ala Pro Glu Tyr Lys Phe Pro Ser Ala Val Ile Asp Ser Phe Asp Ala
        115                 120                 125

ACT AAT TGG GTT TAT AAC AAT TTA GAT AAA TTT GAT GGA AAG ATG GGA
Thr Asn Trp Val Tyr Asn Asn Leu Asp Lys Phe Asp Gly Lys Met Gly
    130                 135                 140

GTT GCG ATT GCG GGA GAT AGT GCT GGA GGA AAT TTG GCA GCG GTT GTA
Val Ala Ile Ala Gly Asp Ser Ale Gly Gly Asn Leu Ala Ala Val Val
145                 150                 155                 160

GCT CTT CTT TCA AAG GGT AAA ATT AAT TTG AAG TAT CAA ATA CTG GTT
Ala Leu Leu Ser Lys Gly Lys Ile Asn Leu Lys Tyr Gln Ile Leu Val
                165                 170                 175

TAC CCA GCG GTA AGT TTA GAT AAC GTT TCA AGA TCC ATG ATA GAG TAC
Tyr Pro Ala Val Ser Leu Asp Asn Val Ser Arg Ser Met Ile Glu Tyr
            180                 185                 190

TCT GAT GGG TTC TTC CTT ACC AGA GAG CAT ATA GAG TGG TTC GGT TCT
Ser Asp Gly Phe Phe Leu Thr Arg Glu His Ile Glu Trp Phe Gly Ser
        195                 200                 205

CAA TAC TTA CGA AGC CCT GCA GAT TTG CTA GAC TTT AGG TTC TCT CCA
Gln Tyr Leu Arg Ser Pro Ala Asp Leu Leu Asp Phe Arg Phe Ser Pro
    210                 215                 220
```

```
ATT CTG GCG CAA GAT TTC AAC GGA TTA CCT CCA GCC TTG ATA ATA ACA
Ile Leu Ala Gln Asp Phe Asn Gly Leu Pro Pro Ala Leu Ile Ile Thr
225                 230                 235                 240

GCA GAA TAC GAT CCA CTA AGG GAT CAA GGA GAA GCG TAT GCA AAT AAA
Ala Glu Tyr Asp Pro Leu Arg Asp Gln Gly Glu Ala Tyr Ala Asn Lys
                245                 250                 255

CTA CTA CAA GCT GGA GTC TCA GTT ACT AGT GTG AGA TTT AAC AAC GTT
Leu Leu Gln Ala Gly Val Ser Val Thr Ser Val Arg Phe Asn Asn Val
            260                 265                 270

ATA CAC GGA TTC CTC TCA TTC TTT CCG TTG ATG GAG CAA GGA AGA GAT
Ile His Gly Phe Leu Ser Phe Phe Pro Leu Met Glu Gln Gly Arg Asp
        275                 280                 285

GCT ATA GGT CTG ATA GGG TCT GTG TTA AGA CGA GTA TTT TAT GAT AAA
Ala Ile Gly Leu Ile Gly Ser Val Leu Arg Arg Val Phe Tyr Asp Lys
    290                 295                 300

ATT TAA
Ile
305
```

Figure 11
LA11.1 Esterase es23

```
ATG AAG GTT AAA CAC GTT ATT GTT TTA CAT GGC TTA TAT ATG TCT GGC TTG GTG ATG CGC    60
Met Lys Val Lys His Val Ile Val Leu His Gly Leu Tyr Met Ser Gly Leu Val Met Arg    20

CCG TTA TGT TCG CGT CTA GAA GAG TCG GGG GTT AAA GTT TTA AAC TTA ACC TAC AAT ACT   120
Pro Leu Cys Ser Arg Leu Glu Glu Ser Gly Val Lys Val Leu Asn Leu Thr Tyr Asn Thr    40

CGA GAC CCT AAT CGA GAT GCT ATT TTT ACG CAA ATA GAT GAG TTT ATT AGC AAT GAG CCT   180
Arg Asp Pro Asn Arg Asp Ala Ile Phe Thr Gln Ile Asp Glu Phe Ile Ser Asn Glu Pro    60

TCT GCT TTA GTG TGT CAC TCT ATG GGG GGC TTA GTT GCT CGC GCC TAT TTA GAG GCA AAC   240
Ser Ala Leu Val Cys His Ser Met Gly Gly Leu Val Ala Arg Ala Tyr Leu Glu Ala Asn    80

TCA GCG CCA AGT CAT CAT GTT GAA AAG GTA ATC ACC TTA GGA ACG CCA CAT ACT GGC AGC   300
Ser Ala Pro Ser His His Val Glu Lys Val Ile Thr Leu Gly Thr Pro His Thr Gly Ser   100

CAT ATT GCT GAA AAA ATG CAG CAA AAA GGG TTC GAG CTA TTA TTA AAA AAT AGC GTT GAG   360
His Ile Ala Glu Lys Met Gln Gln Lys Gly Phe Glu Leu Leu Leu Lys Asn Ser Val Glu   120

TTT TTA CTC TCT AAG AAT GGT GAT TGG CCT TTT AAA GCC AAG CTA TAT AGC ATT GCC GGC   420
Phe Leu Leu Ser Lys Asn Gly Asp Trp Pro Phe Lys Ala Lys Leu Tyr Ser Ile Ala Gly   140

GAC TTA CCG ATT GGC TTA ATG CCA CTC ATT GTA AAA GGC AGC CGC TCT GAT GGC ACT GTA   480
Asp Leu Pro Ile Gly Leu Met Pro Leu Ile Val Lys Gly Ser Arg Ser Asp Gly Thr Val   160

TTG CTA GAT GAA ACC AAG CTA AAG GGT ATG GCT GAA CAC AAG GTG TTT CAT TTA AGC CAT   540
Leu Leu Asp Glu Thr Lys Leu Lys Gly Met Ala Glu His Lys Val Phe His Leu Ser His   180

ACA AGT ATG ATT TAC TCT CGC CAA GTC GTT AAT TAT ATT CTT GAG CGC TTG AAC GAG GAC   600
Thr Ser Met Ile Tyr Ser Arg Gln Val Val Asn Tyr Ile Leu Glu Arg Leu Asn Glu Asp   200

ATT TA                                                                            605
Ile                                                                               201
```

Figure 12

Whale Mat Sample 11.801 Esterase es9

```
ATG ATA AAA AAC TTC GAC AGA GAA AAT TCT AGC TTA GTA CTG TCC GGT GGT GGT GCT CTG    60
Met Ile Lys Asn Phe Asp Arg Glu Asn Ser Ser Leu Val Leu Ser Gly Gly Gly Ala Leu    20

GGT ATT GCT CAC TTG GGT GTA CTG CAT GAC CTT GAA AAA CAA AAT ATT GTA CCA AAT GAA   120
Gly Ile Ala His Leu Gly Val Leu His Asp Leu Glu Lys Gln Asn Ile Val Pro Asn Glu    40

ATT GTT GGT ACA AGT ATG GGT GGT ATC ATT GGT GCA TCT ATG GCT ATC GGG ATG AAA GAG   180
Ile Val Gly Thr Ser Met Gly Gly Ile Ile Gly Ala Ser Met Ala Ile Gly Met Lys Glu    60

AAA GAA ATA CTC GAA GAA ATC AAA AAC TTT TCC AAT GTC TTC AAC TGG ATA AAA TTC TCT   240
Lys Glu Ile Leu Glu Glu Ile Lys Asn Phe Ser Asn Val Phe Asn Trp Ile Lys Phe Ser    80

TTT TCC GGT AAT TCT GTT GTC GAT AAC GAG AAG ATC GCT AAG ATA TTT GAT ACT CTT TTT   300
Phe Ser Gly Asn Ser Val Val Asp Asn Glu Lys Ile Ala Lys Ile Phe Asp Thr Leu Phe   100

AAA GAC AGA AAG ATG ACA GAT ACG GTG ATC CCT CTT AAA CTC ATC GCT ACA AAC TTA CAT   360
Lys Asp Arg Lys Met Thr Asp Thr Val Ile Pro Leu Lys Leu Ile Ala Thr Asn Leu His   120

AAT GGA CAT AAA AAA GTA TTT ACT GCT TCG GAT GAT GTA CTG ATC AAA GAT GCA ATA CTC   420
Asn Gly His Lys Lys Val Phe Thr Ala Ser Asp Asp Val Leu Ile Lys Asp Ala Ile Leu   140

TCA ACA ATG GCA ATA CCC GGT GTA TTT GAA GAA CAT ATT ATT GAT GGT GAA ACC TAT GGC   480
Ser Thr Met Ala Ile Pro Gly Val Phe Glu Glu His Ile Ile Asp Gly Glu Thr Tyr Gly   160

GAC GGT TTT CTT TGT GAA AAC CTT GGT GTG AAT GAG GCA ACA TTC AAT GAT GTT TTA GCT   540
Asp Gly Phe Leu Cys Glu Asn Leu Gly Val Asn Glu Ala Thr Phe Asn Asp Val Leu Ala   180

GTA GAT GTC ATG GGT GAG AAC TCT TTT GAA AAA GCA ATG CCG GAC AAC TTC TTT AAA ACA   600
Val Asp Val Met Gly Glu Asn Ser Phe Glu Lys Ala Met Pro Asp Asn Phe Phe Lys Thr   200

TCA AAT GTT TTA GAA ATG TTT GAA AAA TCA ATG CGA CTT TTT ATT TAC AAC CAG ACA CAG   660
Ser Asn Val Leu Glu Met Phe Glu Lys Ser Met Arg Leu Phe Ile Tyr Asn Gln Thr Gln   220

ACA CAT ATT AAA AAT GCA AAT AAA AAT ATT TAT CTT ATT GAA CCC GTT ACC AAA GAG TAT   720
Thr His Ile Lys Asn Ala Asn Lys Asn Ile Tyr Leu Ile Glu Pro Val Thr Lys Glu Tyr   240

AAA ACA TTT CAA TTT CAT AAA CAT AAA GAG ATA CGT GCT TTA GGC TTG GGT TTA CTG TG    779
Lys Thr Phe Gln Phe His Lys His Lys Glu Ile Arg Ala Leu Gly Leu Gly Leu Leu       259
```

Figure 13

Metallosphaera Prunae Ron 12/2 Esterase 23mcl

```
ATG CCC CTA CAT CCA AAG GTA AAG AAA TTA CTT TCC CAG CTA CCT CCC CAG GAC TTC TCC    60
Met Pro Leu His Pro Lys Val Lys Lys Leu Leu Ser Gln Leu Pro Pro Gln Asp Phe Ser    20

AGA AAC GTG CAG GAC CTG AGG AAG GCC TGG GAT TTA CCC TTC TCA GGG AGG AGG GAG ACC   120
Arg Asn Val Gln Asp Leu Arg Lys Ala Trp Asp Leu Pro Phe Ser Gly Arg Arg Glu Thr    40

CTG AAG AGG GTT GAG GAC CTT GAG ATA CCC ACT AGG GAC GCA CGA ATC AGG GCC AGG GTC   180
Leu Lys Arg Val Glu Asp Leu Glu Ile Pro Thr Arg Asp Ala Arg Ile Arg Ala Arg Val    60

TAC ACC CCC TCA AGT AAG GAA AAC TTA CCC GTC CTT GTT TAC TAT CAC GGC GGT GGC TTC   240
Tyr Thr Pro Ser Ser Lys Glu Asn Leu Pro Val Leu Val Tyr Tyr His Gly Gly Gly Phe    80

GTG TTC GGT AGC GTT GAC AGC TAC GAC GGC CTC GCA TCC CTT ATT GCC AAG GAA TCT GGG   300
Val Phe Gly Ser Val Asp Ser Tyr Asp Gly Leu Ala Ser Leu Ile Ala Lys Glu Ser Gly   100

ATT GCG GTT ATC TCC GTG GAG TAT AGG CTC GCC CCT GAG CAC AAG TTC CCC ACC GCA GTC   360
Ile Ala Val Ile Ser Val Glu Tyr Arg Leu Ala Pro Glu His Lys Phe Pro Thr Ala Val   120

AAC GAC TCG TGG GAT GCG CTT CTC TGG ATC GCG GAG AAC GGA GGC AAG CTG GGG CTC GAC   420
Asn Asp Ser Trp Asp Ala Leu Leu Trp Ile Ala Glu Asn Gly Gly Lys Leu Gly Leu Asp   140

ACC TCG AGA CTT GCC GTG GCT GGG GAT AGT GCT GGA GGA AAC CTG TCT GCC GTG GTG TCC   480
Thr Ser Arg Leu Ala Val Ala Gly Asp Ser Ala Gly Gly Asn Leu Ser Ala Val Val Ser   160

CTC CTG GAC AGG GAC CAG GGT AAG GGA CTG GTT AGT TAT CAG GTC CTA ATC TAC CCA GCA   540
Leu Leu Asp Arg Asp Gln Gly Lys Gly Leu Val Ser Tyr Gln Val Leu Ile Tyr Pro Ala   180

GTG AAC ATG GTC GAT AAC TCC CCA TCC GTC AGG GAG TAC GGC GAG GGA TAC TTC CTC ACC   600
Val Asn Met Val Asp Asn Ser Pro Ser Val Arg Glu Tyr Gly Glu Gly Tyr Phe Leu Thr   200

AGG TCC ATG ATG AAC TGG TTC GGG ACC ATG TAC TTC TCC TCT GGA AGG GAA GCG GTA TCC   660
Arg Ser Met Met Asn Trp Phe Gly Thr Met Tyr Phe Ser Ser Gly Arg Glu Ala Val Ser   220

CCC TAC GCC TCT CCA GCC TTG GCT GAC CTA CAT AAC CTC CCA CCC TCA CTG GTG ATC ACT   720
Pro Tyr Ala Ser Pro Ala Leu Ala Asp Leu His Asn Leu Pro Pro Ser Leu Val Ile Thr   240

GCA GAG TAT GAT CCC CTA AGG GAT CAG GGA GAG ACC TAC TCT CAC TCC CTA AAC GAG GCT   780
Ala Glu Tyr Asp Pro Leu Arg Asp Gln Gly Glu Thr Tyr Ser His Ser Leu Asn Glu Ala   260

GGA AAC GTA TCA ACC TTG GTT AGA TAT CAA GGA ATG ATT CAC GGC TTC CTG TCC TTC TAC   840
Gly Asn Val Ser Thr Leu Val Arg Tyr Gln Gly Met Ile His Gly Phe Leu Ser Phe Tyr   280

GAG TGG ATA ACT GCC GGT AAA CTA GCC ATT CAC CAC ATT GCT GGG GTT CTG AGA TCT GTC   900
Glu Trp Ile Thr Ala Gly Lys Leu Ala Ile His His Ile Ala Gly Val Leu Arg Ser Val   300

CTT TA                                                                             905
Leu                                                                                301
```

Figure 14

Thermotoga neapolitana 5068 Esterase 56mc4

```
GTG GCC TTC TTC GAT ATG CCC CTT GAG GAA CTG AAA AAG TAC CGG CCT GAA AGG TAC GAG    60
Val Ala Phe Phe Asp Met Pro Leu Glu Glu Leu Lys Lys Tyr Arg Pro Glu Arg Tyr Glu    20

GAG AAA GAT TTC GAT GAG TTC TGG AGG GAA ACA CTT AAA GAA AGC GAA GGA TTC CCT CTG   120
Glu Lys Asp Phe Asp Glu Phe Trp Arg Glu Thr Leu Lys Glu Ser Glu Gly Phe Pro Leu    40

GAT CCC GTC TTT GAA AAG GTG GAC TTT CAT CTC AAA ACG GTT GAA ACG TAC GAT GTT ACT   180
Asp Pro Val Phe Glu Lys Val Asp Phe His Leu Lys Thr Val Glu Thr Tyr Asp Val Thr    60

TTC TCT GGA TAC AGG GGG CAG AGA ATA AAG GGC TGG CTT CTT GTT CCG AAG TTG GCG GAA   240
Phe Ser Gly Tyr Arg Gly Gln Arg Ile Lys Gly Trp Leu Leu Val Pro Lys Leu Ala Glu    80

GAA AAG CTT CCA TGC GTC GTG CAG TAC ATA GGT TAC AAT GGT GGA AGG GGT TTT CCA CAC   300
Glu Lys Leu Pro Cys Val Val Gln Tyr Ile Gly Tyr Asn Gly Gly Arg Gly Phe Pro His   100

GAC TGG CTG TTC TGG CCG TCA ATG GGT TAC ATC TGT TTT GTC ATG GAC ACC AGG GGG CAG   360
Asp Trp Leu Phe Trp Pro Ser Met Gly Tyr Ile Cys Phe Val Met Asp Thr Arg Gly Gln   120

GGA AGC GGC TGG ATG AAG GGA GAC ACA CCG GAT TAC CCT GAG GGT CCA GTC GAT CCA CAG   420
Gly Ser Gly Trp Met Lys Gly Asp Thr Pro Asp Tyr Pro Glu Gly Pro Val Asp Pro Gln   140

TAC CCC GGA TTC ATG ACG AGG GGC ATT CTG GAT CCG GGA ACC TAT TAC TAC AGG CGA GTC   480
Tyr Pro Gly Phe Met Thr Arg Gly Ile Leu Asp Pro Gly Thr Tyr Tyr Tyr Arg Arg Val   160

TTC GTG GAT GCG GTC AGG GCG GTG GAA GCA GCC ATT TCC TTC CCG AGA GTG GAT TCC AGG   540
Phe Val Asp Ala Val Arg Ala Val Glu Ala Ala Ile Ser Phe Pro Arg Val Asp Ser Arg   180

AAG GTG GTG GTG GCC GGA GGC AGT CAG GGT GGG GGA ATC CCC CTT GCG GTG AGT GCC CTG   600
Lys Val Val Val Ala Gly Gly Ser Gln Gly Gly Gly Ile Pro Leu Ala Val Ser Ala Leu   200

TCG AAC AGG GTG AAG GCT CTG CTC TGC GAT GTG CCG TTT CTG TGC CAC TTC AGA AGG GCC   660
Ser Asn Arg Val Lys Ala Leu Leu Cys Asp Val Pro Phe Leu Cys His Phe Arg Arg Ala   220

GTG CAA CTT GTC GAC ACA CAC CCA TAC GTG GAG ATC ACC AAC TTC CTC AAA ACC CAC AGG   720
Val Gln Leu Val Asp Thr His Pro Tyr Val Glu Ile Thr Asn Phe Leu Lys Thr His Arg   240

GAC AAA GAG GAG ATT GTT TTC AGA ACA CTT TCC TAC TTC GAT GGT GTG AAC TTT GCA GCA   780
Asp Lys Glu Glu Ile Val Phe Arg Thr Leu Ser Tyr Phe Asp Gly Val Asn Phe Ala Ala   260

AGG GCA AAG GTG CCC GCC CTG TTT TCC GTT GGG CTC ATG GAC ACC ATC TGT CCT CCC TCG   840
Arg Ala Lys Val Pro Ala Leu Phe Ser Val Gly Leu Met Asp Thr Ile Cys Pro Pro Ser   280

ACG GTC TTC GCC GCT TAC AAC CAC TAC GCC GGT CCA AAG GAG ATC AGA ATC TAT CCG TAC   900
Thr Val Phe Ala Ala Tyr Asn His Tyr Ala Gly Pro Lys Glu Ile Arg Ile Tyr Pro Tyr   300

AAC AAC CAC GAA GGT GGA GGT TCT TTC CAG GCA ATT GAG CAG GTG AAA TTC TTG AAG AGA   960
Asn Asn His Glu Gly Gly Gly Ser Phe Gln Ala Ile Glu Gln Val Lys Phe Leu Lys Arg   320

CTA TTT GAG GAA GGC TAG                                                           978
Leu Phe Glu Glu Gly                                                               326
```

Figure 15

Melittangium lichenicola Esterase 77mc1

```
ATG CGC ACC CTC TCC TTC GGT CCG ATG ACC ACA GGG GGA AGC ATT CAC ATG GCG ACC ATG    60
Met Arg Thr Leu Ser Phe Gly Pro Met Thr Thr Gly Gly Ser Ile His Met Ala Thr Met    20

GAC GTG ATG CGC GGG CCG GGG ATG CAG CGG CTG TCA CAG GGC GCC AGG GAG GCC GCG AAC   120
Asp Val Met Arg Gly Pro Gly Met Gln Arg Leu Ser Gln Gly Ala Arg Glu Ala Ala Asn    40

CAC CCC TGG GCG AAG CGA CTG GGC CGC ATG GGC TAC GCG GCC AAG GGC GCC GTG TAC GCC   180
His Pro Trp Ala Lys Arg Leu Gly Arg Met Gly Tyr Ala Ala Lys Gly Ala Val Tyr Ala    60

ATC ATC GGC GTG CTC GCG CTG AAG CTC GCG GCG GGC GAG GGC GGC CGG ACC ACG GAC AGC   240
Ile Ile Gly Val Leu Ala Leu Lys Leu Ala Ala Gly Glu Gly Gly Arg Thr Thr Asp Ser    80

CAC GGC GCG GTG AAC ACC GTG GCG CAC GGG CCC TTC GGC GTC GCG CTG CTG GCG GTG CTG   300
His Gly Ala Val Asn Thr Val Ala His Gly Pro Phe Gly Val Ala Leu Leu Ala Val Leu   100

GTG GTG GGC CTG CTG GGC TAC GTG GTC TGG AGG TTC GCC CAG GCC TTC GTG GAC ACG GAG   360
Val Val Gly Leu Leu Gly Tyr Val Val Trp Arg Phe Ala Gln Ala Phe Val Asp Thr Glu   120

GAC AAG GGC TCC GAC GCG AAG GGA ATC GCC ACG CGC GCC ATG TAC TTC CTC AGC GGC TGC   420
Asp Lys Gly Ser Asp Ala Lys Gly Ile Ala Thr Arg Ala Met Tyr Phe Leu Ser Gly Cys   140

ATC TAC GCG TCG CTG GCC TTC TTC GCC GCG CAG TCC CTG GTG GGC GCC GCG CAC GGC CGG   480
Ile Tyr Ala Ser Leu Ala Phe Phe Ala Ala Gln Ser Leu Val Gly Ala Ala His Gly Arg   160

AGC AAG GGG ACG CAG GGC TGG ACG GCC ACG CTG ATG GAG CAG CCC TTT GGC CGC GTG CTG   540
Ser Lys Gly Thr Gln Gly Trp Thr Ala Thr Leu Met Glu Gln Pro Phe Gly Arg Val Leu   180

GTG GCG CTG GTG GGG CTG GGC ATC GTG GGC TTC GCG CTG AAG CAG TTC CAC ACC GCG TGG   600
Val Ala Leu Val Gly Leu Gly Ile Val Gly Phe Ala Leu Lys Gln Phe His Thr Ala Trp   200

AAG GCG AAG TTC CGG GAG AAG CTC ACC CTC ACC GGA CTG GCT GCC CGG AAG CAG CAC CAC   660
Lys Ala Lys Phe Arg Glu Lys Leu Thr Leu Thr Gly Leu Ala Ala Arg Lys Gln His His   220

ATC GAG CGC ATG TGC CAG TTC GGC ATC GCC GCG CGC GGC GTG GTG TTC GCC GTC ATC GGC   720
Ile Glu Arg Met Cys Gln Phe Gly Ile Ala Ala Arg Gly Val Val Phe Ala Val Ile Gly   240

GGC TTC CTC GTC CGC TCC GCC GTG GAC GCG AAC CCC GGC GAG GCC AAG GGC CTG GGA GAG   780
Gly Phe Leu Val Arg Ser Ala Val Asp Ala Asn Pro Gly Glu Ala Lys Gly Leu Gly Glu   260

GCC CTG GCC GTC GTC GCG AGG CAG CCG TCC GGC GAC GTG CTC CTG GGG GTG GTG GCG GCG   840
Ala Leu Ala Val Val Ala Arg Gln Pro Ser Gly Asp Val Leu Leu Gly Val Val Ala Ala   280

GGC CTG GTG GCC TAC GCC GCC TAC CTG TTC CTC CAG GCG CGC TAC CGC GAA CTC TAG       897
Gly Leu Val Ala Tyr Ala Ala Tyr Leu Phe Leu Gln Ala Arg Tyr Arg Glu Leu End       299
```

Figure 16

Whale Mat Sample 11.801 Esterase es2

```
ATG AGC AAA TTC GCA ATA CTC TGG GCG TTG ATA ACG GCA TAC CTG CCG GAA CCT GTG ATG    60
Met Ser Lys Phe Ala Ile Leu Trp Ala Leu Ile Thr Ala Tyr Leu Pro Glu Pro Val Met    20

AAA CTG GTA TAT TTA GGG CGG CGC GAA ACG CTT GGG GCA CGG ACG CTT GAC GTT AAA GCC   120
Lys Leu Val Tyr Leu Gly Arg Arg Glu Thr Leu Gly Ala Arg Thr Leu Asp Val Lys Ala    40

CAA GCT GTC GGG CGG CTG GCC AAT GCA ACA AGA CCT GTC GGG GTG ATT CCG ACG GTC GAG   180
Gln Ala Val Gly Arg Leu Ala Asn Ala Thr Arg Pro Val Gly Val Ile Pro Thr Val Glu    60

GAA AGC CGG AAG ATG ACG GAT AAA GCC GTT AGC CTT TTT GAT CAG CCC GCC CCC GAA TTA   240
Glu Ser Arg Lys Met Thr Asp Lys Ala Val Ser Leu Phe Asp Gln Pro Ala Pro Glu Leu    80

TTC CGT AAA AAA GAC ATT CAG ATT GAC GGG GCT GAA GGG CCT ATT GAT GCC CGT ATT TAC   300
Phe Arg Lys Lys Asp Ile Gln Ile Asp Gly Ala Glu Gly Pro Ile Asp Ala Arg Ile Tyr   100

AGC GGC CCT GCA AAA CAT CGC CCN CGN CCA ATN CTA GTG TAT TTT CAC GGC GGT GGC TGG   360
Ser Gly Pro Ala Lys His Arg Pro Arg Pro Ile Leu Val Tyr Phe His Gly Gly Gly Trp   120

GTT CAG GGC AAT CTG GAC AGC CAT GAC GGG GTT TGC GGC AAG CTG GCA AAA TGG GCG AAC   420
Val Gln Gly Asn Leu Asp Ser His Asp Gly Val Cys Gly Lys Leu Ala Lys Trp Ala Asn   140

TGC ATT GTT ATC TCG GTC GAT TAT CGT CTA GCG CCC GAA CAC AAA TTT CCT TGT GCG CCG   480
Cys Ile Val Ile Ser Val Asp Tyr Arg Leu Ala Pro Glu His Lys Phe Pro Cys Ala Pro   160

CTT GAT GCG ATT GCG GCC TAT AAA TGG GTG CGC GCC AAC GCA ACA AAC CTT GGC GGC GAT   540
Leu Asp Ala Ile Ala Ala Tyr Lys Trp Val Arg Ala Asn Ala Thr Asn Leu Gly Gly Asp   180

CCT GAA CGT ATC GGC GTT GGC GGC GAT AGC GCA GGG GGC AAT CTT GCC GCC GTT GTC TGC   600
Pro Glu Arg Ile Gly Val Gly Gly Asp Ser Ala Gly Gly Asn Leu Ala Ala Val Val Cys   200

CAA CAA ACC GCC ATG AAC GGC GAG CGC ACA CCA GAT CTG CAA GTC CTG ATC TAT CCG GCG   660
Gln Gln Thr Ala Met Asn Gly Glu Arg Thr Pro Asp Leu Gln Val Leu Ile Tyr Pro Ala   220

CTG GAT GCA CGC ATG ATC TCG ACC TCG ATG GAG GAA TTG CGT GAT GCC TAC ATC TTG CCG   720
Leu Asp Ala Arg Met Ile Ser Thr Ser Met Glu Glu Leu Arg Asp Ala Tyr Ile Leu Pro   240

AAA TCC AGA ATG GAG TAT TTC CTC GGC CTA TAT ACG CGT GGC CCT GAC GAT ATC GAG GAC   780
Lys Ser Arg Met Glu Tyr Phe Leu Gly Leu Tyr Thr Arg Gly Pro Asp Asp Ile Glu Asp   260

CTT AGG ATG TCG CCA ATT CTC AGG GAT ACC GTC GCG GAT CAA CCC CAA GCC TGC ATT GTC   840
Leu Arg Met Ser Pro Ile Leu Arg Asp Thr Val Ala Asp Gln Pro Gln Ala Cys Ile Val   280

ACC TGT GGG TTT GAC CCT GCG CGA CGA CGG GAA CAC CTA CGC CGA ACG CTT AAT TGC CGA   900
Thr Cys Gly Phe Asp Pro Ala Arg Arg Arg Glu His Leu Arg Arg Thr Leu Asn Cys Arg   300

GGG GAT AGA CGT TA                                                                 914
Gly Asp Arg Arg                                                                    304
```

Figure 17

Whale Mat Sample AD3059 Esterase es4

```
GTG AGC ATT CGT CTG CGA CTG TTA AAC TGG TTT TTG AAT ACC TTT GAA AAA CCA AAA CTG    60
Val Ser Ile Arg Leu Arg Leu Leu Asn Trp Phe Leu Asn Thr Phe Glu Lys Pro Lys Leu    20

GCC GCG GCC AAA ACG CCG GAT GAT TTG CGA AAA TCG TTT GAA TTA AAG GCG AGG TTT TTG   120
Ala Ala Ala Lys Thr Pro Asp Asp Leu Arg Lys Ser Phe Glu Leu Lys Ala Arg Phe Leu    40

TTT CCG GCG CCA CGT AAA ACA AGG TTT AGT CAT GAT GTA TTG CAG TCA GGC ATC GGG TCG   180
Phe Pro Ala Pro Arg Lys Thr Arg Phe Ser His Asp Val Leu Gln Ser Gly Ile Gly Ser    60

GTA AAT GCC CAG TGG GCG AAA TCC AAA TCT GCA TCT GAT GAC AGG GTA ATC CTG TAT TTT   240
Val Asn Ala Gln Trp Ala Lys Ser Lys Ser Ala Ser Asp Asp Arg Val Ile Leu Tyr Phe    80

CAT GGG GGA GGG TAT GTT TTT GGG TCA CCA AAA ACG CAC CGT GCA ATG TTG GCG CGC TTG   300
His Gly Gly Gly Tyr Val Phe Gly Ser Pro Lys Thr His Arg Ala Met Leu Ala Arg Leu   100

TCG GCA ATG ACA GGT CTT TCT GCG TGC CTT CCA GAT TAT AGG TTG GCA CCA GAG CAC CCA   360
Ser Ala Met Thr Gly Leu Ser Ala Cys Leu Pro Asp Tyr Arg Leu Ala Pro Glu His Pro   120

TTT CCA GCC GCG ATC GAA GAT GCA GTT TTA TCG TAT AAA TGT TTA CTA GAG CGA GCA ATC   420
Phe Pro Ala Ala Ile Glu Asp Ala Val Leu Ser Tyr Lys Cys Leu Leu Glu Arg Ala Ile   140

GAG CCC CAA AAT ATT ATA CTG GGG GGG GAC AGT GCT GGT GGC GGT TTG GTT CTT GCT TTG   480
Glu Pro Gln Asn Ile Ile Leu Gly Gly Asp Ser Ala Gly Gly Gly Leu Val Leu Ala Leu   160

CTT GCA GAA ATC AAG GCC CAA TCC TTG CCC AAA CCT GCT GGC GTT TTT GCC TTG TCG CCT   540
Leu Ala Glu Ile Lys Ala Gln Ser Leu Pro Lys Pro Ala Gly Val Phe Ala Leu Ser Pro   180

TTG GTT GAT TTA TCA TTT TCG GGC CTT TCG TTT TCT AAA AAT GCC CAA ACC GAT GTG ATG   600
Leu Val Asp Leu Ser Phe Ser Gly Leu Ser Phe Ser Lys Asn Ala Gln Thr Asp Val Met   200

TTG CCC GCA TCA CGG GCT GCG GAT ATG GCG ACC TTG TAT TTG GAT GGG GCC GAT GCA GAT   660
Leu Pro Ala Ser Arg Ala Ala Asp Met Ala Thr Leu Tyr Leu Asp Gly Ala Asp Ala Asp   220

GAT CCA CGT GCA TCG CCG CTG CAG GCG GAT TTT TCT GGC ATG CCG CCT GTA TTT CTG ACA   720
Asp Pro Arg Ala Ser Pro Leu Gln Ala Asp Phe Ser Gly Met Pro Pro Val Phe Leu Thr   240

GCA AGT GAC AGT GAA ATC CTG TTG GAT GAT TGC CTG CGG ATG GCG GAT CAC TTG CGT GCG   780
Ala Ser Asp Ser Glu Ile Leu Leu Asp Asp Cys Leu Arg Met Ala Asp His Leu Arg Ala   260

CAA GGT GTC GTT GTG ACA GAC CGG ATT GTT GAA AAC CAT CCA CAT GTT TGG CAT ATT TTT   840
Gln Gly Val Val Val Thr Asp Arg Ile Val Glu Asn His Pro His Val Trp His Ile Phe   280

CAA CGC CTT CTA CCC GAA GCA GAT CAG GGG CTG CGG GCG ATT GCC GCG TGG ATT AAA CCT   900
Gln Arg Leu Leu Pro Glu Ala Asp Gln Gly Leu Arg Ala Ile Ala Ala Trp Ile Lys Pro   300

CTT TTA TCA GGT TCA AAC GAA AGC TA                                                926
Leu Leu Ser Gly Ser Asn Glu Ser                                                   308
```

Figure 18

Microscilla furvescens Esterase 53ec2

```
ATG CTT ACA TTT AAT GTT TTA TAT GGT ATG ATG AAA CAA AAA CTA GCA GCA ATT CTC ATG    60
Met Leu Thr Phe Asn Val Leu Tyr Gly Met Met Lys Gln Lys Leu Ala Ala Ile Leu Met    20

TTT TTA GGG CTA TCA GCA GCA GAG GCT CAA GAC TGG CCT GAC CTA CAG AAA TAT CGT AGT   120
Phe Leu Gly Leu Ser Ala Ala Glu Ala Gln Asp Trp Pro Asp Leu Gln Lys Tyr Arg Ser    40

GCT AAT AAA GAA GCC AAA TTA CTT CCA AAG GAA AAC CGG AAG GTG GTT TTT ATG GGC AAC   180
Ala Asn Lys Glu Ala Lys Leu Leu Pro Lys Glu Asn Arg Lys Val Val Phe Met Gly Asn    60

TCC ATT ACA GAA GCC TGG ATT AGT CAG CGA CCT GAG TTT TTT AGT GAA AAT GGG TTT ATC   240
Ser Ile Thr Glu Ala Trp Ile Ser Gln Arg Pro Glu Phe Phe Ser Glu Asn Gly Phe Ile    80

GGT CGA GGC ATC AGT GGC CAG ACA ACC CCT CAG ATG TTG TTG AGA TTC CGA CAG GAT GTG   300
Gly Arg Gly Ile Ser Gly Gln Thr Thr Pro Gln Met Leu Leu Arg Phe Arg Gln Asp Val   100

ATA GAC CTG CAG CCA AAG GCT GTA GTG ATA CTA GCT GGT ACC AAT GAC GTA GCT CAA AAT   360
Ile Asp Leu Gln Pro Lys Ala Val Val Ile Leu Ala Gly Thr Asn Asp Val Ala Gln Asn   120

ACC GGG CCG ATG ACC ATT GAG GAA TCG CTT GCT AAC ATT AAG TCT ATG GTG GAG CTG GCG   420
Thr Gly Pro Met Thr Ile Glu Glu Ser Leu Ala Asn Ile Lys Ser Met Val Glu Leu Ala   140

CAA GCC AAT GGG ATC ACG CCT GTT TTG TGT ACC GTG CTG CCT GCA GAT CGT TTC AGC TGG   480
Gln Ala Asn Gly Ile Thr Pro Val Leu Cys Thr Val Leu Pro Ala Asp Arg Phe Ser Trp   160

CGA CCT GAG CTT ACA CCC GCA GAA ACT ATC ATT GCC CTC AAT CAG CTC ATT AAG CAA TAT   540
Arg Pro Glu Leu Thr Pro Ala Glu Thr Ile Ile Ala Leu Asn Gln Leu Ile Lys Gln Tyr   180

GCC GAG GCA CAG GGC CTG GCC CTG GTG GAT TAT CAT GCT GCA CTC ACC AAT AAA GGT GGA   600
Ala Glu Ala Gln Gly Leu Ala Leu Val Asp Tyr His Ala Ala Leu Thr Asn Lys Gly Gly   200

GGA CTT CCG GTG AAA TAC GGA GAA GAT GGT GTG CAT CCA AAT GTA GCA GGC TAT CAG GTG   660
Gly Leu Pro Val Lys Tyr Gly Glu Asp Gly Val His Pro Asn Val Ala Gly Tyr Gln Val   220

ATG GAA AAC ATT GTT TTA CCG GTC ATT TCC AGC GAG TTG GCA AAG CTG AAG TA            713
Met Glu Asn Ile Val Leu Pro Val Ile Ser Ser Glu Leu Ala Lys Leu Lys               237
```

Figure 19

Thermotoga maritima MSB8 Esterase 6ecl

```
ATG GCC TTC TTC GAT TTA CCA CTC GAA GAA CTG AAG AAA TAT CGT CCA GAG CGG TAC GAA   60
Met Ala Phe Phe Asp Leu Pro Leu Glu Glu Leu Lys Lys Tyr Arg Pro Glu Arg Tyr Glu   20

GAG AAA GAC TTC GAT GAG TTC TGG GAA GAG ACA CTC GCA GAG AGC GAA AAG TTC CCC TTA  120
Glu Lys Asp Phe Asp Glu Phe Trp Glu Glu Thr Leu Ala Glu Ser Glu Lys Phe Pro Leu   40

GAC CCC GTC TTC GAG AGG ATG GAG TCT CAC CTC AAA ACA GTC GAA GCG TAC GAT GTC ACC  180
Asp Pro Val Phe Glu Arg Met Glu Ser His Leu Lys Thr Val Glu Ala Tyr Asp Val Thr   60

TTC TCC GGA TAC AGG GGA CAG AGG ATC AAA GGG TGG CTC CTT GTT CCA AAA CTG GAA GAA  240
Phe Ser Gly Tyr Arg Gly Gln Arg Ile Lys Gly Trp Leu Leu Val Pro Lys Leu Glu Glu   80

GAA AAA CTT CCC TGC GTT GTG CAG TAC ATA GGA TAC AAC GGT GGA AGA GGA TTC CCT CAC  300
Glu Lys Leu Pro Cys Val Val Gln Tyr Ile Gly Tyr Asn Gly Gly Arg Gly Phe Pro His  100

GAC TGG CTG TTC TGG CCT TCT ATG GGT TAC ATA TGT TTC GTC ATG GAT ACT CGA GGT CAG  360
Asp Trp Leu Phe Trp Pro Ser Met Gly Tyr Ile Cys Phe Val Met Asp Thr Arg Gly Gln  120

GGA AGC GGC TGG CTG AAA GGA GAC ACA CCG GAT TAC CCT GAG GGT CCC GTT GAC CCT CAG  420
Gly Ser Gly Trp Leu Lys Gly Asp Thr Pro Asp Tyr Pro Glu Gly Pro Val Asp Pro Gln  140

TAT CCA GGA TTC ATG ACA AGA GGA ATA CTG GAT CCC AGA ACT TAC TAC TAC AGA CGA GTC  480
Tyr Pro Gly Phe Met Thr Arg Gly Ile Leu Asp Pro Arg Thr Tyr Tyr Tyr Arg Arg Val  160

TTC ACG GAC GCT GTC AGA GCC GTT GAA GCT GCT GCT TCT TTT CCT CAG GTA GAT CAA GAA  540
Phe Thr Asp Ala Val Arg Ala Val Glu Ala Ala Ala Ser Phe Pro Gln Val Asp Gln Glu  180

AGA ATC GTG ATA GCT GGA GGC AGT CAG GGT GGC GGA ATA GCC CTT GCG GTG AGC GCT CTC  600
Arg Ile Val Ile Ala Gly Gly Ser Gln Gly Gly Gly Ile Ala Leu Ala Val Ser Ala Leu  200

TCA AAG AAA GCA AAG GCT CTT CTG TGC GAT GTG CCG TTT CTG TGT CAC TTC AGA AGA GCA  660
Ser Lys Lys Ala Lys Ala Leu Leu Cys Asp Val Pro Phe Leu Cys His Phe Arg Arg Ala  220

GTA CAG CTT GTG GAT ACG CAT CCA TAC GCG GAG ATC ACG AAC TTT CTA AAG ACC CAC AGA  720
Val Gln Leu Val Asp Thr His Pro Tyr Ala Glu Ile Thr Asn Phe Leu Lys Thr His Arg  240

GAC AAG GAA GAA ATC GTG TTC AGG ACT CTT TCC TAT TTC GAT GGA GTG AAC TTC GCA GCC  780
Asp Lys Glu Glu Ile Val Phe Arg Thr Leu Ser Tyr Phe Asp Gly Val Asn Phe Ala Ala  260

AGA GCG AAG ATC CCT GCG CTG TTT TCT GTG GGT CTC ATG GAC AAC ATT TGT CCT CCT TCA  840
Arg Ala Lys Ile Pro Ala Leu Phe Ser Val Gly Leu Met Asp Asn Ile Cys Pro Pro Ser  280

ACG GTT TTC GCT GCC TAC AAT TAC TAC GCT GGA CCG AAG GAA ATC AGA ATC TAT CCG TAC  900
Thr Val Phe Ala Ala Tyr Asn Tyr Tyr Ala Gly Pro Lys Glu Ile Arg Ile Tyr Pro Tyr  300

AAC AAC CAC GAG GGA GGA GGC TCT TTC CAA GCG GTT GAA CAG GTG AAA TTC TTG AAA AAA  960
Asn Asn His Glu Gly Gly Gly Ser Phe Gln Ala Val Glu Gln Val Lys Phe Leu Lys Lys  320

CTA TTT GAG AAA GGC TAA                                                          978
Leu Phe Glu Lys Gly                                                              326
```

Figure 20

Polyangium brachysporum Esterase 78mcl

```
TTG AAG TAC TTC AAA GCC CGG CTT GCC GGC ATC ACC TTG CTC GGC CTG CTG GCC TGC ACC   60
Leu Lys Tyr Phe Lys Ala Arg Leu Ala Gly Ile Thr Leu Leu Gly Leu Leu Ala Cys Thr   20

TCG GCC TCG GCG CAG ACC GAG CCC ATC GTG TTC GTG CAC GGC TAT TCC GGC AGC GCA TCC  120
Ser Ala Ser Ala Gln Thr Glu Pro Ile Val Phe Val His Gly Tyr Ser Gly Ser Ala Ser   40

AAC TGG GAC ACC ATG CTG GGC CGC TTC CGG TCG AAC GGT TAT GCG TCC GGC TCG CTC TAC  180
Asn Trp Asp Thr Met Leu Gly Arg Phe Arg Ser Asn Gly Tyr Ala Ser Gly Ser Leu Tyr   60

ACC TTC AAC TAC AAC TCG TTG GTC AGC AGC AAC CGC ACC AGC GCC AGC GAG CTG CGC AGC  240
Thr Phe Asn Tyr Asn Ser Leu Val Ser Ser Asn Arg Thr Ser Ala Ser Glu Leu Arg Ser   80

TTC GTC AAC ACC GTG CGT TCG CGC CAC GGC AAC GCC CGC ATC GCG CTG GTC GCC CAC TCC  300
Phe Val Asn Thr Val Arg Ser Arg His Gly Asn Ala Arg Ile Ala Leu Val Ala His Ser  100

AAC GGC GGG CTG GTG TCG CGC TGG TAT CGC GCG GAG CTG GGC GGC GAA ACG GCC ACC CGC  360
Asn Gly Gly Leu Val Ser Arg Trp Tyr Arg Ala Glu Leu Gly Gly Glu Thr Ala Thr Arg  120

CGC TTC GTG ACG CTG GGC ACG CCG CAC CGG GGC ACC ACC TGG GCC TAT GCG TGC TAC AGC  420
Arg Phe Val Thr Leu Gly Thr Pro His Arg Gly Thr Thr Trp Ala Tyr Ala Cys Tyr Ser  140

CCC GCA TGT TTC GAG ATG CGC CCC GGC TCC AGC TTG CTG ACC ACG CTG GGC TCG CGT GCC  480
Pro Ala Cys Phe Glu Met Arg Pro Gly Ser Ser Leu Leu Thr Thr Leu Gly Ser Arg Ala  160

TGC GAC CGC TCG CTG TGG TCG AAC ACC GAC GGC ATC ATC CTG CCG GCG TCC AGC GCG CAG  540
Cys Asp Arg Ser Leu Trp Ser Asn Thr Asp Gly Ile Ile Leu Pro Ala Ser Ser Ala Gln  180

TGT GGT GTC AGC ACG CGC ACT GCC GAC GTC AGC CAT CTC GAC CTG CTG ACC GAC TCT CGC  600
Cys Gly Val Ser Thr Arg Thr Ala Asp Val Ser His Leu Asp Leu Leu Thr Asp Ser Arg  200

GTG TAC ACG CAG TTG CGC ACG CAG TTG CAA TGA GGG TGA CGG TGC ACC GAA CGT GCA CCT  660
Val Tyr Thr Gln Leu Arg Thr Gln Leu Gln End Gly End Arg Cys Thr Glu Arg Ala Pro  220

G                                                                                661
```